DASINYA JOURNAL for engineering and informatics

https://doi.org/10.65542/djei.v2i1.19
Dasinya Journal for Engineering and Informatics. **2026**, 2, 6.

Review

Deep Learning-Based Skin Disease Detection and Classification: A Systematic Literature Review

Zhehat Rebar Abdulqader [1, 2*], Araz Rajab Abrahim [3]

[1] College of Education, University of Zakho, Zakho, Kurdistan Region, Iraq;
[2] Department of Information Technology, Technical College of Informatics, Akre University for Applied Science, Akre, Kurdistan Region, Iraq;
[3] Technical Institute of Duhok, Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq;
* Correspondence: zhehat.abdulqader@uoz.edu.krd

Abstract

Recent advances in deep learning have significantly transformed medical diagnostics, particularly in dermatology. Accurate skin disease detection and classification areessential for effective treatment and improved patient outcomes. This systematic review examines deep learning approaches, including Convolutional Neural Networks (CNNs) and transfer learning, for automated dermatological diagnosis. Public datasets such as HAM10000 and ISIC play a key role in training robust models; however, challenges including dataset imbalance, disease heterogeneity, and overfitting remain. Techniques such as ensemble learning, attention mechanisms, explainable artificial intelligence, data augmentation, hybrid models, and task-specific loss functions have been shown to enhance accuracy, robustness, and interpretability. This study follows a systematic review methodology in accordance with the PRISMA guidelines. The review synthesizes 17 studies published between 2021 and 2024, highlighting the potential of deep learning to support scalable and reliable dermatological diagnostic systems.

Keywords: Deep Learning; Convolutional Neural Networks (CNN); Transfer Learning; Image Processing; Skin Lesions; Image Classification.

1. Introduction

Skin cancer is one of the most prevalent cancers globally, imposing a significant clinical burden on healthcare systems and presenting serious quality-of-life challenges for patients. Skin diseases encompass a wide range of issues, from benign dermatological disorders to rare and complex conditions. These conditions present significant challenges to the field of dermatology due to their intricate appearances and potential health impacts. Early and accurate identification of skin diseases is critical for timely treatment, and recent AI-driven tools have both accelerated and sharpened this process in clinical settings. Traditional diagnostic methods often rely on the dermatologist's judgment, which can be constrained by factors such as availability, accessibility, and the subjective nature of visual examinations, making them prone to errors. Recent advancements in artificial intelligence and deep learning have opened new pathways to automate the detection and classification of skin diseases [1].

Convolutional Neural Networks (CNNs), along with other deep learning methods, have become an effective instrument in medical diagnostics. They help to automate and improve the diagnostic process, which results in better patient outcomes and reduces the workload of medical specialist [2]. The Convolutional Neural Networks are able to learn complex features of images with consideration of large annotated data sets and induce significant improvement on di-agnostic accuracy with minimum human intervention [3]. Other than that, deep learning architecture selection is the key to the success of skin disease detection systems. Different architectures have been studied and combination of

Received: November 29, 2025; Revised: January 14, 2026; Accepted: January 16, 2026; Published: January 25, 2026; Available online: January 25, 2026.

deep learning and transfer learning methods has also increased the efficiency of these systems. This method is especially applicable to medical imaging, where information can be scarce, and it might need re-labeling. By exploiting the acquired features of large datasets, including ImageNet, before refining them on task-specific tasks, e.g. skin diseases classification, such models can greatly decrease training time and increase accuracy by exploiting the acquired features of large datasets [4]. In addition, Ensemble learning techniques involving the use of more than one model have been demonstrated to improve the classification accuracy by alleviating the weaknesses of single models [5].

In addition to the importance of the deep learning model, it is also important to select the dataset. The ability of such models to process skin disease datasets enables creation of strong classifiers to perform well in the generalization of different populations and skin types [6,7]. This diversity of diseases manifestations and types of skin exhibited in this database gives a strong basis on which models can be successful in detecting and differentiating different classes of skin disorders [8].

Although the progress of deep learning in skin diseases diagnosis is promising, there are still a number of problems. The issues of overfitting, large labeled datasets and interpretability of model predictions remain major challenges [9]. In addition, problems like skew of the dataset where some skin diseases are underrepresented may lead to biased predictions in the models [10]. Moreover, the application of these models to the clinical practice also requires a close attention to the ethical considerations, data privacy, and the inclusion of AI systems into the current patterns of work of the healthcare practices [11]. These difficult situations are important aspects that need to be addressed in order to ensure a successful implementation of the deep learning technology in dermatology. There is a need to make sure that these technologies will aid and support, but not substitute the choices made by the healthcare professionals [12].

This review synthesizes recent deep learning methods, CNNs, transfer learning, ensembles, XAI applied to skin disease image datasets. Section 2 presents cover background. The Methodology is presented in Section 3. The results of the review analysis, including the most significant literature, are presented in Section 4. Challenges and Future Work can be found in Section 5. Finally, Section 6 provides a conclusion of the research findings.

2. Background and Theoretical Framework

2.1 Skin Diseases and Their Impact

The fact that skin diseases are visible and prevalent makes them quite challenging to diagnose and thus demand accurate and specialized diagnostic tools [13]. The current developments in deep learning methods have demonstrated the potential to increase the precision of skin disease detection, upon which early treatment and intervention are based. Early diagnosis also translates into improved clinical outcomes of the disease and costs incurred in the healthcare management of advanced diseases are minimized [14]. Additionally, the knowledge of the psychosocial effects of skin diseases can enable healthcare professionals to create holistic care plans that would support both physical and psychological requirements [15].

2.2 Deep Learning in Medical Image Analysis

A breakthrough in (AI) and especially deep learning (CNNs), deep residual networks and transformer architectures has facilitated the classification, diagnosis and segmentation of lesions on the skin with precision and without human intervention. Detection of skin diseases in dermatoscopic image scans and automated detection of histopathological images are just some of the applications that have shown higher diagnostic capabilities than the conventional methods [16]. Recent studies have focused on improving the model robustness and generalizability through learning multi-scale features, attention and with large annotated data sets [17,18].

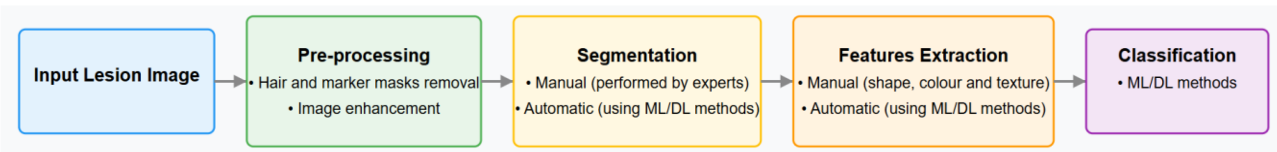

Figure 1. CAD pipeline for skin diseases image analysis.

As shown in Figure 1. typical computer-aided diagnosis (CAD) pipeline of dermoscopic images, the raw input lesion image is the starting point of the pipeline. then is followed by pre-processing steps, which involve hair removal, marker removal and contrast enhancement. This is followed by segmentation that can be done manually by experts or automatically using (ML) or (DL) techniques. Features are extracted, either manually by using ML models or automatically by using DL models. These characteristics are eventually fed into classification model to generate a decision.

2.3 Deep Learning Techniques for Skin Disease Detection and Classification

The latest developments in deep learning methods have been very beneficial in the process of detection and classification of skin diseases. Different methodologies such as CNNs have been utilized to automate the classification process and they have proven to be highly accurate to identify different dermatological conditions. CNNs are enhanced versions of the Artificial Neural Networks (ANNs) which build on the principle of the ANNs by adding another layer of hidden features to the network. An example of this network structure which is deepening is CNNs, as depicted in Figure 2. The effectiveness of hybrid models is shown by such hybrid architectures as U-Net and DenseNets [19], which combine both spatial and contextual features. Moreover, the emphasized models, such as CNNs, employ the methods of transfer learning and ensembles to enhance the classification accuracy. Moreover, data augmentation and pre-trained models have also contributed to performance improvement to a considerable extent [20].

Additionally, research has highlighted the transfer learning and ensemble techniques to increase the performance of classification, especially in the context of multiclass [21]. Combining image processing methods with deep learning systems has also enhanced the accuracy of the diagnosis and provided more effective computer-aided diagnosis systems. These developments have shown promising clinical decision support outcomes in diagnostic errors reduction and helping dermatologists detect skin cancer at its early stages [22].

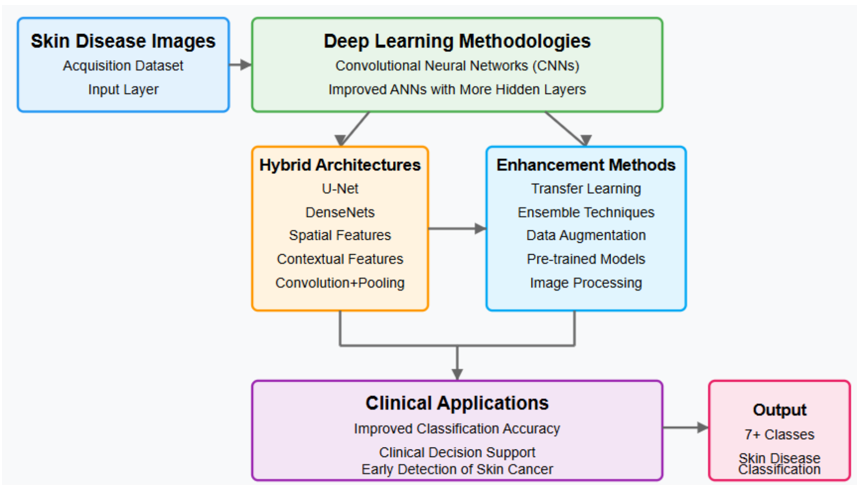

Figure 2. Advanced Deep Learning Techniques for Skin Disease Classification.

Figure 2. illustrates an abstracted DL pipeline to analyze skin diseases, in which a dermoscopic image is first used and a multiclass diagnosis is the final step. CNNs further improved ANNs with extra hidden layers learn fine diseases characteristics, and hybrid models such as U-Net and DenseNets combine space and context information. Transfer learning, ensemble approaches, data augmentation and image-processing pre-steps can also enhance performance so that they are correctly classified into seven and above disease categories. This technique is built into clinical decision support and helps to minimize diagnostic errors and detects skin cancer early.

3.4 Skin Diseases Datasets

2. 4. 1. Dermatoscopic Datasets

Deep dermoscopy is an important non-surgical procedure to determine the high-resolution images of the subsurface skin architecture that greatly improves the early melanoma and other skin diseases detection. With the growing use of AI, ML, and DL in this domain, there is a growing need to have publicly available image datasets.

Such dermoscopic datasets can be used to train and fine-tune the AI, ML, and DL models to guarantee their precise functioning with a wide range of skin types and diseases variations [23]. The most popular public dermoscopic image datasets are summarized in Table. 1, which provides detailed information about each dermoscopic image datasets that are publicly available [24].

Table 1. Skin Disease Datasets Overview.

Dataset Name	Collection Site	Year	No. Disease	Dataset Size
ISIC	**Memorial Sloan Kettering Cancer Center**	**2020**	**5**	**11,108**
BCN20000	**Hospital Clinic Barcelona**	**2019**	**9**	**19,424**
HAM10000	**Medical University of Vienna and skin cancer practice of Cliff Rosendahl in Queensland**	**2018**	**8**	**10,015**
SNU	**University of Edinburgh**	**2018**	**134**	**2,201**
Asan	**Asan Institutional**	**2017**	**12**	**17,125**
PH2	**Dermatology Service of Pedro Hispano Hospital**	**2013**	**3**	**200**

2. 4. 2. HAM-10000 Datasets

The HAM10000 dataset is a very important resource in skin disease detection and classification. This dataset contains 10,015 dermatoscopic images that are divided into seven different types of skin diseases, as they are shown in Table. 2, and it shows how diseases are classified under the HAM10000 dataset.

Table 2. Diversity of HAM-10000 Dataset.

Diseases Name	Category Code	No. of Images	Total Samples (%)
Melanocytic nevi	**NV**	**6,705**	**66.95%**
Melanoma	**MEL**	**1,113**	**11.11%**
Benign keratosis-like lesions	**BKL**	**1,099**	**10.97%**
Basal cell carcinoma	**BCC**	**514**	**5.13%**
Actinic keratoses and intraepithelial carcinoma	**AKIEC**	**327**	**3.27%**
Vascular lesions	**VASC**	**142**	**1.42%**
Dermatofibroma	**DF**	**115**	**1.15%**
Total number of samples		**10,015**	**100.00%**

The dataset has a 644 x 450-pixel image resolution. The dermatoscopic images are of JPEG format. Diversity of the dataset makes the algorithms more robust and they can be generalized to different skin conditions [25], Figure 3. below illustrates the diseases.

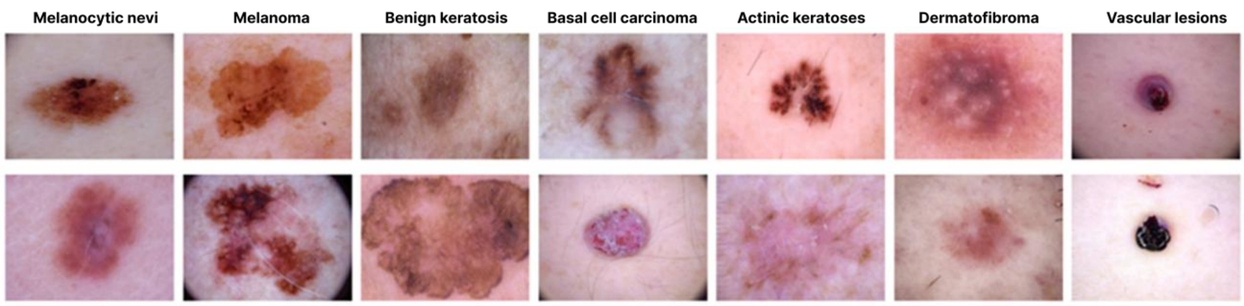

Figure 3. HAM-10000 skin disease categories.

2.5 Explainable AI and Skin Disease Classification

Explainable AI (XAI) techniques offer interpretable outputs, which can be validated and trusted by physicians with AI-based conclusions [26,27]. Moreover, deep learning models tend to make decision-making processes more complicated and therefore become inefficient in high-stakes settings, including healthcare. Thus, to ensure the successful implementation of AI in clinical practice, it is necessary to increase the level of clarity.

In order to elaborate on the intensive study in fashion, a number of methods have been proposed. Explanatory clinical imaging artificial intelligence is inherently human-oriented and is informed by the requirements and concepts of medical staff members [28]. Through these interpretative measures, the clinical subject will be able to increase the credibility and reliability of the adoption of AI in the healthcare sector.

3. Methodology

3.1 Aim and Scope

This study examines and discusses the most recent developments in deep learning methods that can be used to detect and classify skin diseases. It will determine the efficacy of different architectures and approaches such as CNN and transfer learning to overcome such issues as imbalance in the dataset and image diversity. The proposed research will provide the advantages, drawbacks, and prospects of AI-based clinical technology in dermatology, thus enabling more efficient, precise, and convenient clinical practice.

The current review paper summarizes the state-of-the-art in the field of skin disease detection and classification with the use of deep learning, with special attention to the utilization of skin disease datasets, in particular, HAM10000. With the ability of artificial intelligence, such systems can provide quick, precise, and scalable remedies in the detection of skin conditions, which will eventually result in better patient care and patient outcomes. Deep learning deployment in the clinical practice is bound to revolutionize the sphere of dermatology, enabling to realize the early detection and intervention more readily than ever before.

3.2 Search Strategy and Data Collection

This was done using a strategic key word selection process in order to come up with relevant studies. The keywords were selected based on the main ideas of the research in particular, skin diseases, deep learning architectures, and classification tasks and turned into a list of seed terms and controlled vocabulary descriptors. As an example, terms such as "skin lesion," "dermatoscopic image," and "skin disease" were paired with methodological descriptors including "deep learning," "convolutional neural network (CNN)," "transfer learning," "vision transformer," and "explainable AI (XAI)" synonyms and abbreviations were combined using OR operators, while distinct concepts were linked with AND operators (e.g., ("skin diseases" OR "dermatoscopic image") AND ("deep learning" OR "CNN" OR "transfer learning") AND ("classification" OR "detection")).

Based on this search strategy, the queries were executed across four major scientific database, IEEE Xplore, Scopus, ScienceDirect, and MDPI and filtered them to peer-reviewed articles, English-language, full-text articles published between 2021 and 2024. The first results were filtered by relevance and applied to further narrow down our list of keywords by adding new words as they were discovered like "attention mechanism" or "data augmentation" to increase precision without reducing recall. Lastly, inclusion and exclusion criteria were used to bring up the final corpus of studies that form the basis of this review.

3.3 Review Design and PRISMA Compliance

This study was conducted as a systematic literature review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The PRISMA framework was adopted to ensure methodological transparency, reduce selection bias, and provide a reproducible process for identifying, screening, and selecting relevant studies. Accordingly, the review followed four structured phases: identification, screening, eligibility assessment, and final inclusion, as illustrated in the PRISMA flow diagram Figure 4.

A total of 94 records were identified through searches of IEEE Xplore, Scopus, ScienceDirect, and MDPI. After removing 13 duplicates, 81 unique articles remained and were assessed at the screening stage, with all proceeding to full-text eligibility evaluation. During eligibility assessment, 64 studies were excluded due to insufficient results, outdated or non-novel methods, irrelevance to AI-based skin disease classification, or misalignment with the review scope, while 12 articles were unavailable in full text. Consequently, 17 studies satisfied the inclusion criteria and were retained for qualitative synthesis.

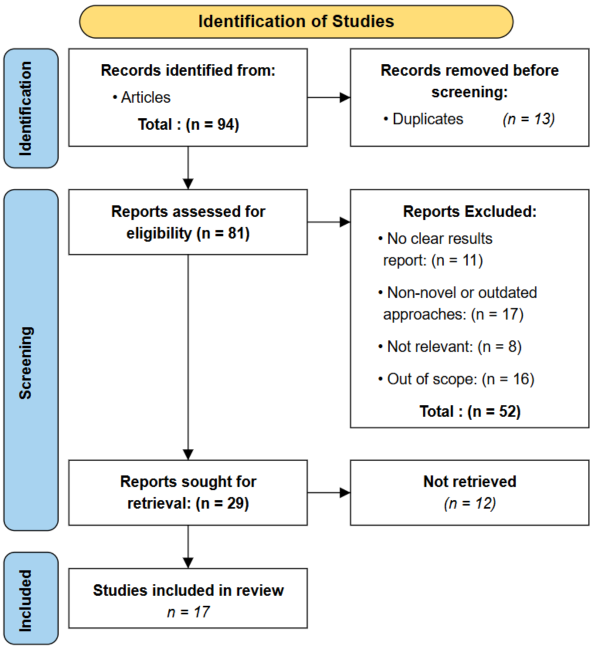

Figure 4. Prisma diagram.

Figure 5 summarizes the methodological distribution of the selected studies, revealing that transfer learning with pre-trained CNNs predominates (41.2%), followed by multi-scale and hybrid architectures and advanced preprocessing techniques (17.6% each). Explainable AI and hand-crafted feature-based approaches each represent 11.8%, indicating the dominance of deep learning while reflecting emerging interest in interpretability and traditional methods.

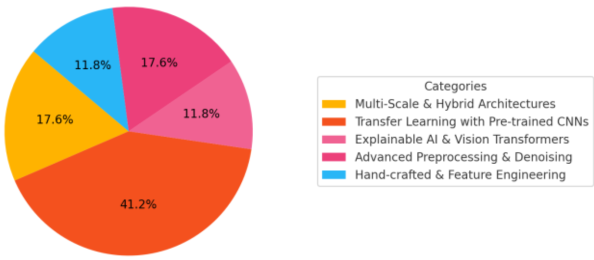

Figure 5. Proportional Distribution of Methodological Clusters in Reviewed Studies.

While Table 3. organizes the individual papers by their respective groups. Together, they highlight where research efforts have been most concentrated and pinpoint opportunities for further innovation.

Table 3. Papers Organized by Methodological Group.

No.	Methodological Group	Authors
1	Multi-Scale & Hybrid Architectures	[29], [30], [31]
2	Transfer Learning with Pre-trained CNNs	[32], [33], [34], [35], [36], [37], [38]
3	Explainable AI & Vision Transformers	[39], [40]
4	Advanced Preprocessing & Denoising	[41], [42], [43]
5	Hand-crafted & Feature Engineering	[44], [45]

3.4 Study Selection Process and Eligibility Criteria (PRISMA-Guided)

The initial database search yielded 94 records. After removing 13 duplicate articles, 81 unique studies remained for screening. During the screening phase, titles and abstracts were reviewed to assess their relevance to AI-based skin disease classification. All records that passed this stage were subsequently evaluated in the eligibility phase through full-text assessment. Study eligibility was determined based on predefined inclusion and exclusion criteria.

Inclusion Criteria: To guarantee that the most pertinent and high-quality studies were included in our review, the following criteria were used:

- Target AI-Driven Classification: Studies need to be focused on machine learning or deep learning methods to detect or classify skin diseases.
- Publication: The article should be published in peer-reviewed journals that are included in one of the following repositories: IEEE Xplore, Scopus, ScienceDirect or MDPI.
- Publication Date: Only articles published in 2021-2024 were put into consideration.
- Language and Accessibility: Full-text articles are offered in English.
- Methodological Transparency: Research should be sufficiently detailed in methodology, that is, the model methods used and the image data used should be described in detail.

Exclusion Criteria: The studies were not included in the review in case they had any of the following criteria:

- The paper is not focused on the application of AI, such as ML or DL, to the identification of skin dis-eases.
- It is a conference abstract, editorial, letter or opinion piece which lacks original research data.
- Studies lacking sufficient experimental or methodological detail.
- The research itself does not use any of the following publicly available dermoscopic datasets: ISIC, HAM10000, PH2, Dermnet.
- Research focused exclusively on lesion segmentation without classification.

Following this selection process, 64 articles were excluded, and 17 studies met all eligibility requirements and were included in the final qualitative synthesis.

3.5 Quality Assessment of Included Studies

To ensure the reliability and scientific rigor of the selected literature, a qualitative quality assessment was conducted for all 17 included studies. Each paper was evaluated based on the following criteria:

- Dataset transparency: Clear description of dataset source, size, and class distribution.
- Methodological clarity: Explicit reporting of model architecture, training strategy, and preprocessing steps.
- Evaluation rigor: Use of appropriate performance metrics (e.g., accuracy, sensitivity) and validation protocols.
- Reproducibility: Sufficient experimental detail to allow replication.
- Clinical relevance: Discussion of applicability, limitations, and potential clinical impact.

Only studies that met the majority of these criteria were retained. Papers with incomplete methodological descriptions or insufficient validation were excluded during the eligibility phase.

4. Results

4.1 Multi-Scale & Hybrid Architectures

Hu et al. in 2024, proposed a multi-scale feature fusion network based on the skilled NETV2 architecture to address challenges in skin wound classification using the HAM10000 and ISIC2019 datasets. The model achieved an accuracy of 94.0%, and 89.8% accuracy on the HAM10000 and ISIC2019 dataset. Such techniques that can be credited to this performance include: restarting against class weighting, label smoothing and class imbalance. The model is effective because of the combination of shallow and deep features that target the wound area, and solutions to problems associated with datasets. Marking the wound area is an essential element of classifying dermoscopic images, though the hair features also play a certain role in the process of the classification [29].

Karthik et al. in 2024, proposed a hybrid deep learning architecture where Swin Transformer was used to extract global features and Dense Group Shuffle Non-Local Attention (DGSNLA) Network was used to extract local features. This model was tested on HAM10000 dataset, it gave a maximum accuracy of 94.21% with recall of 96.25% when both the networks were combined thus increasing the feature representation. Data augmentation methods and focal loss were used to overcome the issue of class imbalance, whereas the use of both local and global features allowed the model to handle both short-and-long-range dependencies successfully [30].

In 2021 Srinivasu et al., suggested a classification model that was built using MobileNet V2 and Long Short-Term Memory (LSTM) was recommended to detect skin disease using dermatoscopic images of the HAM10000 dataset. The given model attained an accuracy of 85.34% with the use of the lightweight MobileNet V2 as a feature extractor and LSTM to process sequential information. It is a state-of-the-art method that integrates MobileNet V2 and the LSTM module to guarantee accurate classification of skin diseases. Also, the fine-tuning and data augmentation were utilized to enhance the efficiency of model training. [31].

Recent innovations in multi-scale and hybrid architectures demonstrate their capabilities in terms of feature fusion (e.g. NETV2-based frameworks), transformer-CNN hybrids of capturing global-local dependencies, as well as light-weight CNNs coupled with sequential modeling. Nevertheless, these studies have serious limitations: they lack sufficient ablation tests to separate architectural efforts with auxiliary methods, lack cross dataset testing (e.g. dermoscopic versus clinical images), and do not address the computational requirements of real-world implementation. Future studies must focus on adaptive fusion (e.g., attention-based weighting), energy efficient transformer-CNN hybrids and intensive benchmarking on a variety of datasets in order to improve clinical relevance.

4.2 Transfer Learning with Pre-trained CNNs

Singh et al. in 2023, tested a pre-trained VGG16 model for skin disease classification by analyzing a massive dataset of 44,000 images from Kaggle. VGG16 model with fine-tuning is an efficient extraction and classification model of images, classifying them as benign or malignant. The model attained a high accuracy 90.1% and recall of 94. 20%. when optimized using hyperparameters, such as the learning rate and epochs. My goal was to use the abilities of VGG16 to recognize and categorize different skin diseases using different images, which demonstrates my strict method. The model is also highly effective with regard to early diagnosis of skin diseases at the various stages [32].

In 2022 Anand et al., proposed a new transfer learning model to skin cancer diagnosis. The ResNet50 architecture was further improved by adding a flatten layer, two dense layers with Leaky ReLU activation and a final dense layer with sigmoid activation. In addition, randomness and augmentation of the dataset were implemented to determine the stability of the model with regard to data augmentation. The ResNet50 model trained on Adam optimizer and 128 batch size reached an accuracy of 90% with sensitivity of 74.42 and the training images were only augmented to enhance the training accuracy of the model, virtually doubling the number of training images [33].

In 2023 Bhargavi et al., presented a model on skin diseases classification using the HAM-10000 dataset, using InceptionResNetV2, InceptionV3, MobileNetV2, and EfficientNetB0. The authors applied data augmentation method to counter the impact of the class imbalance and used pre-trained models to ease the burden of extracting features of images. With the use of predictions and fine-tuning of the model layers, the proposed model was able to obtain a higher accuracy of 81.3% with recall of 80.1% which was higher than the performance of each of the individual models. This method helps in precise and accurate diagnosis of seven types of cancer with the use of wound images. [34].

Rangaswamy et al. in 2024, comparative analysis of skin disease classification with the InceptionV3 and VGG16, CNN models were made on a dataset of 17,214 images of 13 distinct disease categories. InceptionV3 was the most accurate in terms of training with an accuracy of 80.88% compared to VGG16 which had an accuracy of 74.17%. Normalization, flipping and rotation have been used as pre-processing techniques that increase the variability of data. Preprocessing was also done to deal with the issue of class imbalance using shear range and zoom range, which also helped the models to make strong predictions of different diseases on the skin [35].

Jain et al. in 2021, presented six transfer learning models such as Xception, to classify seven types of cancer in the skin based on the HAM10000 dataset. They equalized the classes by replicating the images and applying transformation operations like rotation and zooming, the Xception model had an accuracy of 90.48% and a recall of 89.57% [36].

In 2023 Inthiyaz et al., proposed an automated skin disease detector based on CNN was suggested on the Xiangya-Derm dataset, the largest set of clinical skin images. The researchers pre-processed the pictures by resizing the pictures and used a Softmax classifier to classify the pictures into four classes of eczema, melanoma, psoriasis and healthy skin. The model had an accuracy of 87.42%. It was also optimized based on data augmentation and an optimized architecture. The features were also picked out by counting the features of all the images by downsizing the image to a single measure, which underlines how feature consistency is important in classification [37].

Zhi et al. in 2024, suggested a multiclassification model which is derived through Inception-v2 network and focal loss to analyze dermoscopy images. The model used data augmentation, hair removal and Grad-CAM heat maps to enhance image preprocessing and interpretation. The model was based on the dataset of ISIC 2019 and the accuracy of this model was estimated to be 89.04% and recall of 90.15%. The findings indicate that focal loss can improve the performance of the model, especially in case the dataset is unevenly distributed [38].

The effectiveness of fine-tuning and class imbalance correction methods, including focal loss and data augmentation, are demonstrated with the help of the transfer learning methods involving the VGG16, Inception, and Xception architectures. One of the enduring weaknesses of these approaches is that they are based on curated dermoscopic datasets, and it limits their extrapolation to clinical contexts, where imaging conditions are heterogeneous. Moreover, simple approaches to augmentation do not necessarily use sophisticated data generation algorithms, e.g., GANs. The future studies must be based on the creation of domain adaptation models, multi-modal data (clinical and dermoscopic) and exploration of modern loss functions, including distribution-aware margins, to achieve greater robustness in a variety of imaging settings.

4.3 Explainable AI & Vision Transformers

Hosny et al. in 2024, a novel explainable deep inherent learning architecture proposed to classify multi-class skin diseases was suggested, with the use of a CNN having 54 layers. This was a method of combining both inherent learning and XAI methods to correctly recognize and categorize seven different types of skin diseases. Using the ISIC 2018 database, the model displayed a remarkable accuracy of 92.89% and sensitivity of 58.57% on the basis of the improvements in the information flow of layers, better visualization of the features, and the sensitivity of the map to the occlusion. The learning approach inherent in the proposed model was able to predict several different types of disease and give visual understanding regarding explainability, and as such, it was less susceptible to error compared to traditional shallow networks [39].

In 2023 Arshed et al., The multi-class classification of skin cancer with a Vision Transformer (ViT) model was suggested, and it takes advantage of the feature of ViT to improve the attention drawn to the important parts of a picture by its self-attention mechanism. This model was contrasted to 11 CNN models that applied the fine-tuning and data augmentation strategies on the HAM10000 dataset to solve the issue of class imbalance. ViT-based model also set an accuracy of 92.14% and recall of 92.14%, which was optimized using a mixture of fine-tuning and training. The transformer self-attention mechanism improves the recognition of valuable features and reduces the effects of noise, which shows the strength of this methodology [40].

ViTs and explainable convolutional neural networks are self-attention based and occlusion sensitivity mapping-based vision transformers that provide state-of-the-art accuracy. Nonetheless, the models face the following issues regarding computational efficiency: ViTs demand large amounts of data, and deep CNNs are complex. Also, their explainability, especially in the clinician aligned measures, has not been quantitatively confirmed. Future research directions involve creating hybrid ViT-CNN models to be more efficient, pretraining on large-scale medical images, and creating standardized assessment methods of XAI to make technical explainability more consistent with clinical trust.

4.4 Advanced Preprocessing & Denoising

Kavitha et al. in 2024, proposed CNN-based skin cancer detection system utilizes the ISIC dataset. The authors performed image preprocessing by removing hair and noise to enhance image quality, followed by the application of and ResNet50 models for classification. This method achieved a performance of 91.32% accuracy and recall of 78.15%, employed techniques has an enhancement, of deep learning architecture, and augmentation. This was achieved success of this approach implementing attributed to the implementation of training it on multiple across the convolutional outcome classified nine different successfully of skin cancer [41].

Gururaj et al. in 2023, discussed skin cancer classification using a deep learning model with the HAM10000 dataset. The researchers applied data preprocessing techniques, including the Dull Razor method for noise removal and segmentation using encoder-decoder models. They utilized DenseNet169 and ResNet50 architectures for training, achieving an accuracy of 91.2% and recall of 69.6% through under-sampling using DenseNet169 model and optimized training over multiple epochs for both techniques [42].

In 2024 Pandey et al., a method for skin cancer detection was proposed that combines non-local means (NLM) denoising, sparse dictionary learning, and CNNs. The HAM10000 and ISIC2019 datasets were preprocessed using NLM to enhance image quality. After applying data augmentation and sparse dictionary learning, the trained CNN model achieved accuracies of 85.61% for the HAM10000 dataset and 81.23% for the ISIC2019 dataset. Denoising significantly improved image quality and model performance, particularly by reducing residual noise, which facilitated clearer pattern recognition [43].

The importance of removing noise and eliminating artifacts as part of improving classification performance is emphasized in studies that focus on preprocessing methods, including hair removal and non-local denoising. Nevertheless, preprocessing pipelines are often biased with regard to datasets, e.g. inefficient hair removal in different skin types, and they are usually not based to measure the isolated effect of these processes on model improvements. Future re-search ought to be directed at incorporating adaptive denoising techniques, e.g. learnable filters, into end-to-end frame-works, and making sure that diagnostically important features are not lost in the preprocessing. Also, the cross-domain reliability will be necessary through benchmarking with self-supervised denoising techniques.

4.5 Hand-crafted & Feature Engineering

Kumar et al. in 2024, suggested to identify multiclass skin diseases, based on new hand-crafted features of spatial, spectrogram and cepstrum-domain features. The model has been tested on the HAM10000 and DermNet datasets and yielded 89.71% and 88.57% accuracies respectively with 89.24% and 88.28% recall respectively. These outcomes were achieved by optimizing features, data augmentation and hyperparameter tuning. The concatenated features utilize both spatial and spectral information, which allows deriving more detailed information out of difficult sets of data, which further supports the usefulness of the method [44].

A et al. in 2024, A gradual end-to-end model of skin diseases classification was suggested based on the S-MobileNet CNN model and the HAM10000 dataset. Gaussian filtering was used to perform segmentation of the data and modified SFTA was used to extract features. The S-MobileNet is a lightweight architecture that uses activation functions and compression of the intermediate layer to optimize the performance of the architecture with an accuracy of 89.71% and recall of 89.24%. S-MobileNet CNN architecture was optimized to produce low-latency results. The findings were confirmed by 80:20 training and testing split [45].

Hybrid models that blend hand-designed spatial-spectral inputs with lightweight CNNs are competitive in their accuracy because they use domain-specific feature engineering. Yet, these methods tend to be ineffective in keeping up with the changing pattern of diseases and have not been yet verified in a wide variety of individuals, including different skin tones and atypical subtypes. Future research should integrate hand-crafted features with self-supervised pretraining, employ Neural Architecture Search (NAS) for automated feature optimization, and prioritize the use of inclusive datasets to ensure equitable diagnostic performance across different populations.

Table 4. Summary of Recent Studies in the Field of Skin Diseases.

Authors, Year	Model	Technique	Dataset	Dataset Size	Acc (%)	Sens (%)	Description
[29] Hu et al., 2024	Multi-scale NETV2	Class weighting, label smoothing, resampling	HAM10000 ISIC-2019	10,015 33,569	94.0 89.8	91.7 N/A	Multi-scale feature fusion network integrating shallow and deep features. Focused on lesion regions and handled hair interference.
[30] Karthik et al., 2024	Swin Transformer, DGSNLA	Data augmentation, focal loss function, feature fusion	HAM10000	10,015	94.21	96.25	Combined global and local feature extraction using hybrid Swin Transformer and DGSNLA networks.
[31] Srinivasu et al., 2021	MobileNet V2, LSTM	Data augmentation, lightweight architecture	HAM10000	10,015	85.34	N/A	Combined MobileNet V2 for feature extraction and LSTM for sequence handling in dermatoscopic.

[32] Singh et al., 2023	VGG16	Transfer learning, fine-tuning, hyperparameter optimization.	Open-repository dataset	44,000	90.1	94. 20	Leveraged VGG16 to classify images as benign or malignant. Focused on learning rate and epochs for optimization.
[33] Anand et al., 2022	ResNet50	Data augmentation, LeakyReLU activation, Adam optimizer	HAM10000	10,015	90	74.42	Modified ResNet50 architecture with additional layers and data augmentation to improve performance.
[34] Bhargavi et al., 2023	InceptionResNetV2, InceptionV3, MobileNetV2, EfficientNetB0	Data augmentation, ensemble learning	HAM10000	10,015	81.3	80.1	Utilized pre-trained models for feature extraction and combined predictions for improved accuracy.
[35] Rangaswamy et al., 2024	InceptionV3, VGG16	Normalization, flips, rotations	Roboflow + Kaggle "Skin Melanomas"	17,214	80.88	N/A	A comparative study of different image processing treatments on the nature of the convolutional neural networks for the study so that new approaches for a robust prediction.
[36] Jain et al., 2021	Xception	Transfer learning, data augmentation	HAM10000	10,015	90.48	89.57 (recall)	Denoising and sparse dictionary learning were combined to improve image quality and categorization.
[37] Inthiyaz et al., 2023	CNN with Softmax	Image resizing, Softmax classifier, data augmentation	Xiangya-Derm	150,223	87.42	N/A	Pre-processed clinical dataset for consistency and optimized CNN architecture.
[38] Zhi et al., 2024	Inception-v2	Focal loss, Grad-CAM, hair removal	ISIC 2019	25,332	89.04	90.15 (recall)	To enhance model functioning and interpretability, focal loss and Grad-CAM heat maps were used for imbalanced datasets.
[39] Hosny et al., 2024	54-layer CNN	Inherent learning, XAI techniques, occlusion sensitivity mapping	ISIC 2018	15,414	92.89	58.57	Novel CNN approach with layer-by-layer improvements and explainability for seven skin diseases types.

[40] Arshed et al., 2023	Vision Transformer	Fine-tuning, data augmentation	HAM10000	10,015	92.14	92.14 (recall)	To resolve class imbalance and enhance categorization, ViT was used in conjunction with self-attention and fine-tuning strategies.
[41] Kavitha et al., 2024	CNN-based	Hair removal, noise removal, data augmentation	ISIC	2,357	91.32	78.15	Preprocessing and CNN-based architecture optimized for feature extraction and diverse classification tasks.
[42] Gururaj et al., 2023	DenseNet169, under sampling	Dull Razor method, encoder-decoder models, under-sampling	HAM10000	10,015	91.2	69.6	Images to remove noise and trained models with optimized techniques over multiple epochs.
[43] Pandey et al., 2024	Sparse dictionary based CNN	Sparse dictionary learning, data augmentation	HAM10000 ISIC2019	10,015 25,332	85.61 81.23	N/A	Combined denoising with sparse dictionary learning to enhance classification and image quality.
[44] Kumar et al., 2024	1-D Multiheaded CNN	Spatial, spectrogram, and cepstrum-domain feature integration	HAM10000 Dermnet	10,015 +25,000	89.71 88.57	89.24 88.28	Novel methodology combining spatial and spectral features with augmented data and hyperparameter tuning.
[45] A et al., 2024	S-MobileNet, Gaussian filtering	Mish activation, SFTA feature extraction, layer compression	HAM10000	10,015	89.71	89.24	Developed S-MobileNet with lightweight architecture and Gaussian filtering for segmentation.

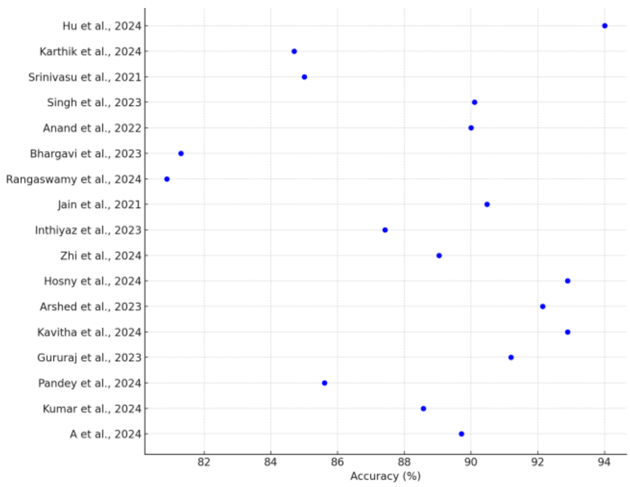

Figure 6. Forest Plot of accuracy across different studies.

Figure 6. synthesizes the estimates of accuracy of seventeen recent studies on the topic of skin diseases classification and shows that there is not only methodological diversity but also variance in performance. The most efficient is reported in [29], with its multi-scale EfficientNetV2 achieving an accuracy of 94.0, and the least one is reported in [35] with 80.9% accuracy. These disparities are an indication of differences in preprocessing methods, feature-fusion approaches, and dataset composition. It is important to note that transformer-based systems such as Swin Transformer + DGSNLA ([30]; 84.7%) and Vision Transformer ([40]; 92.1]) achieve mid and high-performance, indicating that attention mechanisms can be useful in solving complex and multi-class problems. Existing CNN backbones, like VGG16 ([32]; 90.1%), Incep-tionResNetV2 ensembles ([34]; 81.3%), and lightweight networks (e.g., MobileNet; [45]; 89.7%]) fall between 80 and 90 percent, whereas lightweight networks (e.g., MobileNet; [45]; 89.7%]) are promising to use in resource-constrained deployment. Performance differences also indicate dataset scale, of which big clinical repositories ([37]; 150k images; 87.4%) are compared to moderate open-access sets (Ham10000, ~10k images). In general, the plot motivates that incorporating innovative preprocessing methods (hair removal and denoising), attention or hybrid networks, and balanced augmentation strategies are associated with increased accuracy, which will inform future studies to integrate these aspects to achieve effective dermatoscopic analysis.

5. Challenges and Future Research Directions

5.1 Challenges

- Class Imbalance and Diagnostic Risk : Publicly available dermoscopic datasets such as HAM10000 and ISIC are inherently imbalanced, with benign diseases (e.g., melanocytic nevi) dominating the sample distribution while clinically critical malignancies remain underrepresented. As illustrated in Figure 7, class imbalance is the most frequently reported challenge in the reviewed literature. In real-world clinical deployment, this imbalance may result in biased classifiers that under-detect rare but aggressive skin cancers, potentially leading to delayed diagnosis and adverse patient outcomes. Although techniques such as data augmentation, focal loss, and resampling partially mitigate this issue in research settings, their effectiveness in routine clinical environments remains limited without true population-level diversity.

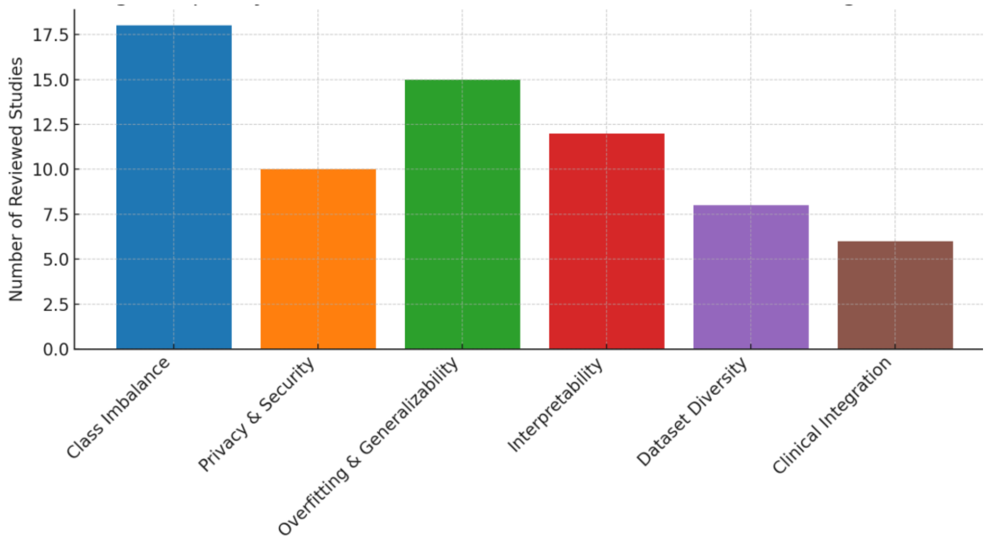

Figure 7. Frequency of challenges discussed in literature.

- Data Privacy, Security, and Regulatory Constraints: The centralized collection and sharing of medical images introduce significant privacy, ethical, and regulatory challenges that directly affect clinical adoption. Strict data protection regulations restrict multi-institutional data pooling, thereby limiting dataset diversity and hindering model generalization. From a clinical perspective, failure to address data governance, consent, and security concerns can prevent regulatory approval and restrict the deployment of AI-assisted diagnostic systems in hospital environments.

- Model Overfitting and Generalizability: Many high-performing deep learning models demonstrate excellent accuracy on curated dermoscopic datasets but experience substantial performance degradation when applied to real-world clinical images acquired using different devices, lighting conditions, and acquisition protocols. This performance gap between experimental and real-world settings is depicted in Figure 8, which contrasts model accuracy in research environments versus clinical scenarios. Such lack of robustness poses a major barrier to clinical translation, as unreliable predictions in routine practice could undermine diagnostic confidence and patient safety.

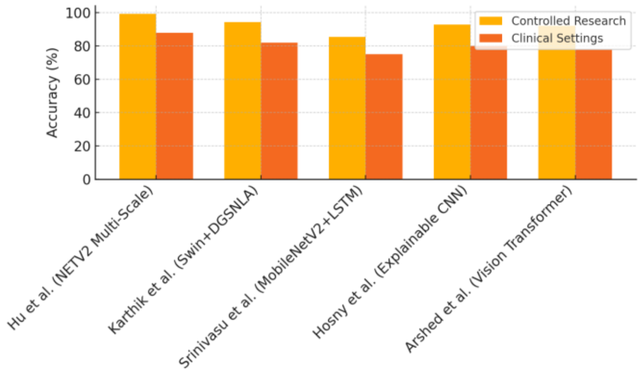

Figure 8. Model Accuracy vs Clinical Settings.

- Interpretability and Clinical Trust: Despite recent advances in explainable AI (XAI), most deep learning models remain black-box systems from a clinician's perspective. While visualization techniques such as Grad-CAM and occlusion sensitivity maps offer some interpretability, they are rarely validated against dermatologist reasoning or clinical decision-making processes. In high-stakes medical environments, the absence of clinically meaningful explanations reduces physician trust and limits the acceptance of AI-based decision support tools.

- Limited Dataset Diversity and Health Equity Concerns: The majority of reviewed studies rely on a narrow set of publicly available datasets, primarily HAM10000 and ISIC, as shown in Figure 9, which compares the frequency of dataset usage across the literature. These datasets insufficiently represent darker skin tones, pediatric populations, and rare dermatological conditions. As a result, models trained on such data may exhibit biased performance, potentially exacerbating healthcare disparities when deployed in clinical practice.

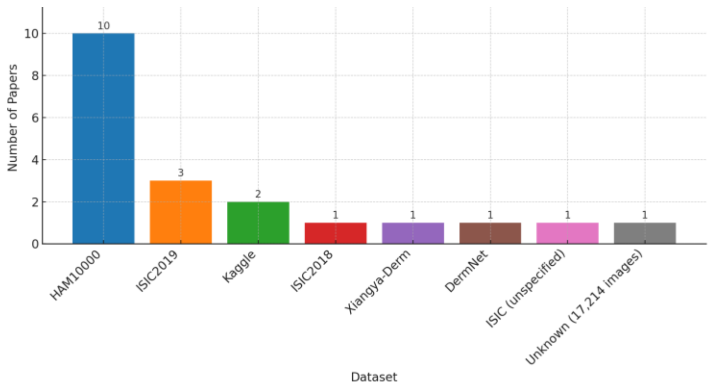

Figure 9. Publicly dataset available.

- Integration into Clinical Workflows: Even when technically accurate, many AI models fail to account for practical deployment constraints such as interoperability with hospital systems, workflow compatibility, and usability for clinicians. Without seamless integration into existing standards (e.g., DICOM, HL7-FHIR) and minimal disruption to clinical routines, deep learning models are unlikely to be adopted in real-world dermatology practice.

5.2 Future Research Directions

To enable effective clinical translation, future research should prioritize the development of class-balanced learning strategies using generative models such as GANs and diffusion models, combined with domain-adaptive loss functions. Privacy-preserving learning paradigms, including federated and split learning, should be adopted to facilitate multi-center collaboration without compromising patient confidentiality. Enhancing model generalizability through self-supervised pretraining, cross-domain validation, and lightweight CNN–Transformer hybrid architectures will be critical for real-world reliability.

Furthermore, standardized and clinically validated explainability frameworks must be developed to align AI decision-making with dermatologist interpretation. Expanding dataset diversity through multi-ethnic, multi-institutional image repositories will be essential to ensure fair and unbiased diagnostic performance. Finally, early collaboration with clinicians, regulatory bodies, and healthcare institutions is necessary to address usability, interoperability, and compliance requirements, thereby accelerating the safe deployment of deep learning-based skin disease diagnostic systems.

In general, Table 5 underlines the key issues related to deep-learning-based skin-disease detection, and classification, as well as the further research directions that should be implemented to resolve these problems. Such challenges as class imbalance, privacy protection, robustness of models, interpretability, and clinical integration are still fundamental barriers to the translation of these systems out of research environments into clinical practice.

Table 5. Challenges and Corresponding Future Directions.

Challenge	Future Directions	References
Class Imbalance and Diagnostic Risk	**Generative models, domain adaptation, class-balanced loss functions.**	**[29]; [30]; [38]; [40].**
Data Privacy, Security, and Regulatory Constraints	**Federated/split learning, differential privacy, secure multi-party computation.**	**[30]; [43].**
Model Overfitting and Generalizability	**Lightweight CNN-Transformer hybrids, self-supervised pretraining, cross-domain validation.**	**[32]; [34]; [36]; [35].**
Interpretability and Clinical Trust	**Diagnostic-concordance metrics, attention gating, prototype learning.**	**[38]; [39].**
Limited Dataset Diversity and Health Equity Concerns	**Multi-ethnic datasets, rare pathology inclusion, style-transfer augmentation.**	**[37]; [43]; [44].**
Integration into Clinical Workflows	**Usability trials, interoperability, regulatory engagement.**	**[31]; [33]; [45].**

6. Conclusion

This review article summarizes the major progress in the deep learning-based approaches to automation of the skin disease detection and classification. It emphasizes significant advances that have been made by using CNNs, hybrid models, transfer learning, and ensemble learning algorithms, which have resulted in high accuracy on large datasets including HAM10000, ISIC and PH2. Other challenges that are discussed in the article are class imbalance, data scarcity, and overfitting. Moreover, explainable AI methods, such as Grad-CAM and LIME, have enhanced the interpretability of models, which is essential to make the method applicable in reality. Nevertheless, there are a few challenges that have remained such as the demand of lightweight, streamlined architectures that can be deployed in real-time, high noise tolerance, and high generalizability in varied datasets. The further studies should focus on the need to develop effective, privacy-saving models and assure credible per-performance in any clinical environment. The issues about the diversity of datasets, model strength, and data security will be critical to the successful implementation of AI-based diagnostic instruments into practice. The combination of deep learning and dermatology is a new frontier in enhancing the early diagnosis, treatment outcome, and patient recovery and it can transform the dermatological care across the globe.

Funding: This research did not receive any specific funding from public, commercial, or non-profit organizations.

Data Availability Statement: HAM10000 is an open-access resource hosted on Harvard dataset [46] (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T).

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical Approval: This study did not involve any experiments on humans or animals and did not require ethical approval. All data used in this research were obtained from publicly available international datasets and used in accordance with the terms and conditions stated by the dataset providers.

References

1. Mushtaq, S.; Singh, O. A Deep Learning Based Architecture for Multi-Class Skin Cancer Classification. *Multimed. Tools Appl.* **2024**, doi:10.1007/s11042-024-19817-1.
2. Naeem, A.; Anees, T.; Fiza, M.; Naqvi, R.A.; Lee, S.W. SCDNet: A Deep Learning-Based Framework for the Multiclassification of Skin Cancer Using Dermoscopy Images. *Sensors* **2022**, *22*, doi:10.3390/s22155652.
3. Almutairi, A.; Khan, R.U. Image-Based Classical Features and Machine Learning Analysis of Skin Cancer Instances. *Appl. Sci.* **2023**, *13*, doi:10.3390/app13137712.
4. Remya, S.; Anjali, T.; Sugumaran, V. A Novel Transfer Learning Framework for Multimodal Skin Lesion Analysis. *IEEE Access* **2024**, *12*, 50738–50754, doi:10.1109/ACCESS.2024.3385340.
5. Pennisi, A.; Bloisi, D.D.; Suriani, V.; Nardi, D.; Facchiano, A.; Giampetruzzi, A.R. Skin Lesion Area Segmentation Using Attention Squeeze U-Net for Embedded Devices. *J. Digit. Imaging* **2022**, *35*, 1217–1230, doi:10.1007/s10278-022-00634-7.
6. Balasundaram, A.; Shaik, A.; Alroy, B.R.; Singh, A.; Shivaprakash, S.J. Genetic Algorithm Optimized Stacking Approach to Skin Disease Detection. *IEEE Access* **2024**, *12*, 88950–88962, doi:10.1109/ACCESS.2024.3412791.
7. Al-Saedi, D.K.A.; Savaş, S. Skin Lesion Classification by Weighted Ensemble Deep Learning. *Iran J. Comput. Sci.* **2024**, *7*, 785–800, doi:10.1007/s42044-024-00210-y.
8. Ji, Z.; Wang, X.; Liu, C.; Wang, Z.; Yuan, N.; Ganchev, I. EFAM-Net: A Multi-Class Skin Lesion Classification Model Utilizing Enhanced Feature Fusion and Attention Mechanisms. *IEEE Access* **2024**, 143029–143041, doi:10.1109/ACCESS.2024.3468612.
9. Darapaneni, N.; Sahni, B.; Paduri, A.R.; Jain, S.; Mohamed, S.; Banerjee, A.; Chakrabarti, S. Early Detection of Skin Cancer - Solution for Identifying and Defining Skin Cancers Using AI. In Proceedings of the 2022 Interdisciplinary Research in Technology and Management (IRTM); 2022; pp. 1–6.
10. Alam, T.M.; Shaukat, K.; Khan, W.A.; Hameed, I.A.; Almuqren, L.A.; Raza, M.A.; Aslam, M.; Luo, S. An Efficient Deep Learning-Based Skin Cancer Classifier for an Imbalanced Dataset. *Diagnostics* **2022**, *12*, doi:10.3390/diagnostics12092115.
11. Mittal, R.; Jeribi, F.; Martin, R.J.; Malik, V.; Menachery, S.J.; Singh, J. DermCDSM: Clinical Decision Support Model for Dermatosis Using Systematic Approaches of Machine Learning and Deep Learning. *IEEE Access* **2024**, *12*, 47319–47337, doi:10.1109/ACCESS.2024.3373539.
12. Noronha, S.S.; Mehta, M.A.; Garg, D.; Kotecha, K.; Abraham, A. Deep Learning-Based Dermatological Condition Detection: A Systematic Review With Recent Methods, Datasets, Challenges, and Future Directions. *IEEE Access* **2023**, *11*, 140348–140381, doi:10.1109/ACCESS.2023.3339635.
13. Stratigos, A.J.; Richard, M.A.; Dessinioti, C.; Paul, C.; Nijsten, T.; Gisondi, P.; Salavastru, C.; Taieb, C.; Trakatelli, M.; Puig, L.; et al. The Prevalence of Skin Diseases in Greece, Impact on Quality of Life and Stigmatization: A Population-Based Survey Study. *JEADV Clin. Pract.* **2024**, *3*, 591–599, doi:10.1002/jvc2.287.
14. Hammad, M.; Pławiak, P.; Elaffendi, M.; El-latif, A.A.A.; Latif, A.A.A. Psoriasis Skin Detection. *Sensors* **2023**, *23*, 7295, doi:doi.org/ 10.3390/s23167295.
15. Jiang, Z.; Gu, X.; Chen, D.; Zhang, M.; Xu, C. Deep Learning-Assisted Multispectral Imaging for Early Screening of Skin Diseases. *Photodiagnosis Photodyn. Ther.* **2024**, *48*, 104292, doi:10.1016/j.pdpdt.2024.104292.
16. Ayaz, A.; Boonstoppel, R.; Lorenz, C.; Weese, J.; Pluim, J.; Breeuwer, M. Effective Deep-Learning Brain MRI Super Resolution

Using Simulated Training Data. *Comput. Biol. Med.* **2024**, *183*, 109301, doi:10.1016/j.compbiomed.2024.109301.
17. Rebar, Z.; Mohsin Abdulazeez, A. Deep and Machine Learning Algorithms for Diagnosing Brain Cancer and Tumors. *Indones. J. Comput. Sci.* **2024**, *13*, doi:10.33022/ijcs.v13i3.4028.
18. Farhan, N.M.; Setiaji, B. Indonesian Journal of Computer Science. *Indones. J. Comput. Sci.* **2023**, *12*, 284–301, doi:10.33022/ijcs.v13i3.4028.
19. Mounica, M.K.V.S. GAN Based Multi-Class Skin Disease Classification: Deep Learning Approach. *Int. J. Res. Appl. Sci. Eng. Technol.* **2024**, *12*, 137–142, doi:10.22214/ijraset.2024.61366.
20. Paramanandam, K.; Kanagavalli, R. A Review on Deep Learning Techniques for Saliency Detection. *Lect. Notes Networks Syst.* **2023**, *400*, 279–289, doi:10.1007/978-981-19-0095-2_29.
21. Tetengi, A.; Abdullahi, M.; Florentin, A.; Kana, D.; Tukur, M.; Hayatu, I. Categorical Classi Fi Cation of Skin Cancer Using a Weighted Ensemble of Transfer Learning with Test Time Augmentation. *Data Sci. Manag.* **2024**, *8*, 174–184, doi:10.1016/j.dsm.2024.10.002.
22. Ibrahim, S.; Amin, K.; Alkanhel, R.I.; Abdallah, H.A.; Ibrahim, M. Soft Attention Based Efficientnetv2b3 Model for Skin Cancer's Disease Classification Using Dermoscopy Images. *IEEE Access* **2024**, *12*, 161283–161295, doi:10.1109/ACCESS.2024.3486153.
23. Grignaffini, F.; Barbuto, F.; Piazzo, L.; Troiano, M.; Simeoni, P.; Mangini, F.; Pellacani, G.; Cantisani, C.; Frezza, F. Machine Learning Approaches for Skin Cancer Classification from Dermoscopic Images: A Systematic Review. *Algorithms* **2022**, *15*, 1–30, doi:10.3390/a15110438.
24. Debelee, T.G. Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review. *Diagnostics* **2023**, *13*, doi:10.3390/diagnostics13193147.
25. Nirupama; Virupakshappa MobileNet-V2: An Enhanced Skin Disease Classification by Attention and Multi-Scale Features. *J. Imaging Informatics Med.* **2024**, *38*, 1734–1754, doi:10.1007/s10278-024-01271-y.
26. Abdulqader, Z.R.; Abdulqader, D.M.; Ahmed, O.M.; Ismael, H.R.; Ahmed, S.H.; Haji, L. Responsible AI Development for Sustainable Enterprises: A Review of Integrating Ethical AI with IoT and Enterprise Systems. *J. Inf. Technol. Informatics* **2024**, *3*, 129–156.
27. Rasel, M.A.; Abdul Kareem, S.; Kwan, Z.; Yong, S.S.; Obaidellah, U. Bluish Veil Detection and Lesion Classification Using Custom Deep Learnable Layers with Explainable Artificial Intelligence (XAI). *Comput. Biol. Med.* **2024**, *178*, 108758, doi:https://doi.org/10.1016/j.compbiomed.2024.108758.
28. Tulsani, V.; Sahatiya, P.; Parmar, J.; Parmar, J. XAI Applications in Medical Imaging: A Survey of Methods and Challenges. *Int. J. Recent Innov. Trends Comput. Commun.* **2023**, *11*, 181–186, doi:10.17762/ijritcc.v11i9.8332.
29. Hu, Z.; Mei, W.; Chen, H.; Hou, W. Multi-Scale Feature Fusion and Class Weight Loss for Skin Lesion Classification. *Comput. Biol. Med.* **2024**, *176*, 108594, doi:10.1016/j.compbiomed.2024.108594.
30. Karthik, R.; Menaka, R.; Atre, S.; Cho, J.; Easwaramoorthy, S.V. A Hybrid Deep Learning Approach for Skin Cancer Classification Using Swin Transformer and Dense Group Shuffle Non-Local Attention Network. *IEEE Access* **2024**, *12*, 158040–158051, doi:10.1109/ACCESS.2024.3485507.
31. Srinivasu, P.N.; Sivasai, J.G.; Ijaz, M.F.; Bhoi, A.K.; Kim, W.; Kang, J.J. Networks with MobileNet V2 and LSTM. **2021**, *21*, 2852, doi:10.3390/s23167295.
32. Singh, G.; Guleria, K.; Sharma, S. A Transfer Learning-Based Pre-Trained VGG16 Model for Skin Disease Classification. *2023 IEEE 3rd Mysore Sub Sect. Int. Conf. MysuruCon 2023* **2023**, 1–6, doi:10.1109/MysuruCon59703.2023.10396942.
33. Anand, V.; Gupta, S.; Nayak, S.R.; Koundal, D.; Prakash, D.; Verma, K.D. An Automated Deep Learning Models for Classification of Skin Disease Using Dermoscopy Images: A Comprehensive Study. *Multimed. Tools Appl.* **2022**, *81*, 37379–37401, doi:10.1007/s11042-021-11628-y.
34. Bhargavi, M.; Renugadevi, R.; Sivabalan, S.; Phani, P.; Ganesh, J.; Bhanu, K. Ensemble Learning for Skin Lesion Classification: A Robust Approach for Improved Diagnostic Accuracy (ELSLC). *3rd Int. Conf. Innov. Mech. Ind. Appl. ICIMIA 2023 - Proc.*

2023, 390–395, doi:10.1109/ICIMIA60377.2023.10425888.

35. Rangaswamy, S.; Tantry, S.S.; Lal, T.S. Skin Disease Classification Using Deep Learning. **2024**, doi:10.1007/s40009-024-01523-z.
36. Jain, S.; Singhania, U.; Tripathy, B.; Nasr, E.A.; Aboudaif, M.K. Deep Learning-Based Transfer Learning for Classification of Skin Cancer. *Sensors* **2021**, *21*, 16, doi:https://doi.org/10.3390/s21238142.
37. Inthiyaz, S.; Altahan, B.R.; Ahammad, S.H.; Rajesh, V.; Kalangi, R.R.; Smirani, L.K.; Hossain, M.A.; Rashed, A.N.Z. Skin Disease Detection Using Deep Learning. *Adv. Eng. Softw.* **2023**, *175*, 103361, doi:10.1016/j.advengsoft.2022.103361.
38. Zhi, S.; Li, Z.; Yang, X.; Sun, K.; Wang, J. A Multiclassification Model for Skin Diseases Using Dermatoscopy Images with Inception-V2. *Appl. Sci.* **2024**, *14*, 10197, doi:https://doi.org/10.3390/app142210197.
39. Hosny, K.M.; Said, W.; Elmezain, M.; Kassem, M.A. Explainable Deep Inherent Learning for Multi-Classes Skin Lesion Classification. *Appl. Soft Comput.* **2024**, *159*, 111624, doi:10.1016/j.asoc.2024.111624.
40. Arshed, M.A.; Mumtaz, S.; Ibrahim, M.; Ahmed, S.; Tahir, M.; Shafi, M. Multi-Class Skin Cancer Classification Using Vision Transformer Networks and Convolutional Neural Network-Based Pre-Trained Models. *Inf.* **2023**, *14*, doi:10.3390/info14070415.
41. Kavitha, C.; Priyanka, S.; Kumar, M.P.; Kusuma, V. Skin Cancer Detection and Classification Using Deep Learning Techniques. *Procedia Comput. Sci.* **2024**, *235*, 2793–2802, doi:10.1016/j.procs.2024.04.264.
42. Gururaj, H.L.; Manju, N.; Nagarjun, A.; Manjunath Aradhya, V.N.; Flammini, F. DeepSkin: A Deep Learning Approach for Skin Cancer Classification. *IEEE Access* **2023**, *11*, 50205–50214, doi:10.1109/ACCESS.2023.3274848.
43. Pandey, A.; Teja, M.S.; Sahare, P.; Kamble, V.; Parate, M. Skin Cancer Classification Using Non - Local Means Denoising and Sparse Dictionary Learning Based CNN. *J. Electr. Syst. Inf. Technol.* **2024**, doi:10.1186/s43067-024-00162-0.
44. Kumar, A.; Vishwakarma, A.; Bajaj, V.; Mishra, S. Novel Mixed Domain Hand-Crafted Features for Skin Disease Recognition Using Multiheaded CNN. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–13, doi:10.1109/TIM.2024.3370772.
45. A, R.S.; Chamola, V.; Hussain, Z.; Albalwy, F.; Hussain, A. A Novel End-to-End Deep Convolutional Neural Network Based Skin Lesion Classification Framework. *Expert Syst. Appl.* **2024**, *246*, 123056, doi:10.1016/j.eswa.2023.123056.
46. Tschandl, P.; Rosendahl, C.; Kittler, H. Data Descriptor: The HAM10000 Dataset, a Large Collection of Multi-Source Dermatoscopic Images of Common Pigmented Skin Lesions. *Sci. Data* **2018**, *5*, 1–10, doi:10.1038/sdata.2018.161.