Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 11, No. 3, 2025

Research on Automatic Recognition and Auxiliary Diagnosis of Artificial Intelligence in Skin Diseases

Ping Li [1, a], Manyi Zhuang [2, b], Jun Tang [3, c], Jian Huang [4, *]

[1] School of Humanities and Management, Youjiang Medical University for Nationalities, Baise Guangxi, 533000, China
[2] School of Public Health, Youjiang Medical University for Nationalities, Baise Guangxi, 533000, China
[3] School of Clinical Medicine, Youjiang Medical University for Nationalities, Baise Guangxi, 533000, China
[4] School of Humanities and Management, Youjiang Medical University for Nationalities, Baise Guangxi, 533000, China

*** Corresponding author:** Jian Huang (Email: 00842@ymun.edu.cn), [a] lipingyyds@zohomail.cn, [b] ElaineChong@outlook.com, [c] ksjbshd@outlook.com

Abstract: Artificial intelligence (AI) has made significant strides in skin disease diagnosis, demonstrating impressive potential in the classification and detection of conditions such as skin cancer. However, several challenges remain that hinder the full clinical adoption of AI-driven diagnostic systems. This paper reviews the current advancements and key research areas aimed at optimizing AI models for dermatological applications. Key areas for future research include the development of more advanced deep learning architectures, such as multimodal fusion methods, which integrate dermoscopic images with structured clinical data for more comprehensive diagnostics. Additionally, there is a growing emphasis on creating personalized AI models that account for individual patient characteristics, such as age, gender, and genetics, to improve diagnostic accuracy. The integration of explainable AI (XAI) techniques, auxiliary tasks like lesion segmentation, and multi-center clinical research is essential to ensure transparency, trust, and generalizability across diverse populations. Moreover, ethical concerns related to bias, data privacy, and accountability must be addressed to ensure fairness and transparency in AI systems. This review highlights the importance of interdisciplinary collaboration and proposes future directions to enhance the reliability, scalability, and clinical applicability of AI-driven skin disease diagnosis, ultimately improving patient outcomes and accessibility to care.

Keywords: Artificial Intelligence (AI); Skin Disease Diagnosis; Multimodal Fusion; Personalized Diagnosis; Explainable AI (XAI); Data Augmentation.

1. Introduction

1.1. Automatic Recognition Techniques

1.1.1. Image - based Recognition Methods

Image-based automatic skin disease recognition has evolved significantly, transitioning from traditional algorithms to modern deep learning methods. Early traditional methods, such as thresholding, edge detection, and histogram analysis, used handcrafted features to identify patterns in skin images, but struggled with issues like image variability and complex lesion morphology, leading to limited accuracy [1]. For example, color histogram analysis for melanoma detection showed success in controlled environments but failed in real-world scenarios with varying skin tones and lighting [2]. These limitations prompted the adoption of deep learning techniques. Recent advancements have been driven by Convolutional Neural Networks (CNNs), which automatically extract hierarchical features from images, making them highly effective for skin disease classification. CNNs have shown performance comparable to dermatologists, with a study achieving 91% accuracy in melanoma classification [4]. However, challenges such as the need for large annotated datasets and the interpretability of model decisions remain [5]. Other deep learning architectures, such as ResNet and attention-based models, have also been explored. ResNet addresses issues in very deep networks, while attention models enhance accuracy by focusing on key features of skin lesions [6, 7]. Additionally, integrating multimodal data, like combining skin images with clinical metadata, has been shown to improve classification accuracy, as in the case of MDFNet, which achieved 92.3% accuracy [8]. Despite these advancements, practical challenges persist. High-quality annotated datasets are difficult and expensive to obtain, and deep learning models can underperform when trained on limited or biased data [9]. Interpretability remains a key concern, as the "black box" nature of deep learning models complicates trust and adoption in clinical practice [5]. Furthermore, performance disparities across different skin tones highlight the need for more inclusive datasets [10, 11]. In addition to classification, deep learning models like U-Net have been used for segmentation and localization, improving diagnostic accuracy by providing detailed lesion maps [12]. However, the integration of these models into clinical practice requires careful validation and regulatory approval, with challenges like data privacy and model bias still to be addressed.

1.1.2. Feature Extraction Algorithms

Feature extraction algorithms are crucial in the automatic recognition of skin diseases, determining the quality of input data for recognition models. This section reviews both manual and automatic methods, particularly those using deep learning. Manual Feature Extraction Methods. Manual methods, relying on predefined rules and statistical techniques, have been widely used in traditional image processing. Color features, such as histograms, RGB values, and HSV representations, are often utilized in skin disease diagnosis. A study [6] showed that color features alone could achieve 75% accuracy in distinguishing between benign and malignant lesions. However, these methods are sensitive to lighting conditions and may not generalize well. Texture features, such as Gray Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP), focus on pixel intensity arrangements. LBP, for instance, achieved 82% accuracy in distinguishing melanoma from non-melanoma lesions [13]. Despite their effectiveness, manual methods are limited by the need for domain expertise and an inability to capture complex features. Automatic Feature Extraction in Deep Learning

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized feature extraction by learning hierarchical features from raw data. CNNs, such as VGGNet and ResNet, have outperformed traditional methods. A study [3] found that ResNet-50 achieved 92.3% accuracy in classifying skin diseases. Deep learning's ability to learn intricate patterns like lesion boundaries and textures from large datasets contributes to this superior performance. One of its advantages is handling large-scale datasets, such as the ISIC dataset with over 20,000 dermoscopy images. This exposure enhances the model's generalization capability. Comparison of Manual and Automatic Methods Manual methods offer interpretability and low computational cost but are limited by the need for expert input and sensitivity to noise. In contrast, deep learning methods provide high accuracy and automatic learning but require substantial computational resources and large datasets. Challenges and Future Directions.

Despite their advantages, deep learning methods face challenges like the need for large annotated datasets, which are time-consuming to obtain. Solutions like transfer learning and data augmentation have been proposed. A study [14] showed that fine-tuning pre-trained models on smaller datasets can yield comparable results to training models from scratch. Another challenge is interpretability; while deep learning models are accurate, their decision-making process is often opaque. Recent developments in explainable AI (XAI), such as Grad-CAM, have been used to visualize the areas of an image most influencing the model's decision, improving trust and interpretability [5].

1.1.3. Machine Learning Models for Recognition

Machine learning models have been key in the automatic recognition of skin diseases, particularly with large-scale image datasets. Traditional models, like support vector machines (SVMs) and decision trees, have been widely used due to their simplicity and interpretability. SVMs, effective in high-dimensional data, achieved 82% accuracy in melanoma and benign nevi classification [15], though they rely on manual feature extraction, which can limit performance. Decision trees, known for their transparency, achieved 78% accuracy in diagnosing common skin conditions like eczema and psoriasis [1], but also face challenges in handling complex datasets. In contrast, deep learning models like ResNet and VGG have revolutionized skin disease recognition by automating feature extraction. ResNet, with its residual connections, mitigates the vanishing gradient problem, achieving 92% accuracy in skin disease classification. VGG, known for its simplicity, reached 90% accuracy in a study [16]. These models are trained on large, labeled datasets, enabling them to generalize well. For example, a hybrid deep learning model trained on 10,000 images achieved 94% accuracy [13]. However, deep learning models are resource-intensive and risk overfitting, especially with small datasets. To improve generalization, techniques like data augmentation (e.g., random cropping and flipping) have been used. A study [17] found that augmentation improved accuracy by 5%. Transfer learning, where a pre-trained model is fine-tuned on a specific dataset, has also proven effective. For instance, fine-tuning a pre-trained ResNet on a skin cancer dataset achieved 95% accuracy [18]. These approaches help bridge the gap between model complexity and data availability. Deep learning models are evaluated using metrics such as accuracy, precision, recall, and F1 score. A study on AI for skin cancer diagnosis reported 91% accuracy, with precision and recall of 89% and 93%, respectively. Despite their effectiveness, deep learning models like ResNet and VGG require significant computational power. For example, training ResNet on a high-performance GPU took about 12 hours. Machine learning models, particularly deep learning, offer high diagnostic accuracy and efficiency, with AI-assisted systems shown to reduce misdiagnosis rates by 30% in primary care [19]. However, the adoption of AI in clinical practice raises concerns about transparency and bias. Future research should address these challenges and explore ways to enhance model performance and clinical applicability.

1.2. Auxiliary Diagnosis Approaches

1.2.1. Rule - based Diagnosis Systems

Rule-based diagnosis systems are commonly used in dermatology to mimic expert decision-making by relying on predefined rules derived from clinical experience and medical knowledge. These systems apply criteria like patient symptoms and medical history to suggest diagnoses, e.g., "if a patient has a red, itchy rash and recent allergen exposure, the diagnosis is likely contact dermatitis." Such rules, codified into decision trees or logical statements, offer a structured approach, reducing subjective judgment. A study by the EADV Task Force on contact dermatitis [20] developed specific diagnostic criteria incorporated into rule-based systems, improving diagnostic consistency. These systems are transparent and interpretable, providing clear reasoning for their conclusions, which is valuable in clinical settings. For example, if diagnosing psoriasis, a rule-based system can list specific symptoms and test results that support the diagnosis, helping clinicians verify the result. Rule-based systems can achieve diagnostic accuracy of up to 85% in controlled settings [17], but they face challenges in staying current with evolving medical knowledge and adapting to new data. For example, the ISIC frequently updates its skin cancer diagnostic guidelines, and rule-based systems may become outdated if not regularly updated. Additionally, capturing complex diagnostic patterns is difficult. Dermatological conditions often overlap, making it hard to create accurate rules for distinguishing similar diseases like eczema and psoriasis. To address these limitations, hybrid systems combining rule-based approaches with machine learning techniques have been explored. For instance, integrating rule-based reasoning with CNNs has been shown to improve diagnostic accuracy [16]. In this approach, rule-based systems filter diagnoses, while machine learning models refine them by analyzing complex data patterns. Rule-based systems are used in clinical settings, especially in telemedicine, where they guide non-specialist clinicians through diagnosis [19]. These systems help reduce dermatologists' workload by handling simpler cases, but they still depend on the quality of the rules. As medical knowledge advances, developing more dynamic, adaptable systems that integrate new data and machine learning could improve their effectiveness, making them more robust and suitable for modern dermatology.

1.2.2. Knowledge - Based Expert Systems

Knowledge-based expert systems are crucial in skin disease diagnosis, relying on structured medical knowledge bases. The process involves acquiring, representing, and storing medical knowledge, with acquisition being the most challenging. It requires extracting relevant information from dermatologists and converting it into a machine-readable format, often using ontologies or NLP techniques. For example, the ISIC has developed a database of dermatoscopic images to aid knowledge acquisition. Knowledge is represented using formal structures like rules or semantic networks and stored in databases or knowledge graphs for efficient retrieval. The reasoning engine applies inference techniques to derive conclusions based on input data, using methods like forward- or backward-chaining. Despite their effectiveness, these systems face challenges in knowledge updates and handling conflicts when multiple rules suggest different conclusions. For example, lesions with overlapping characteristics may complicate diagnosis [21]. They also rely on predefined rules, which limits adaptability to novel or ambiguous cases. Hybrid systems combining expert knowledge and machine learning have improved performance [9]. Data-driven methods represent a shift in diagnostic approaches by leveraging large clinical datasets to build predictive models, with CNNs and random forests commonly used for skin disease classification. These models excel at extracting features from images, such as texture and color, and can handle large, high-dimensional datasets [22]. However, data scarcity, especially for rare diseases, remains a challenge, and data augmentation techniques can help mitigate this. The size and quality of datasets directly affect model performance, with larger datasets leading to better generalization [25][26]. Data-driven models can generalize across different populations, but they also face ethical concerns, such as bias in unrepresentative datasets. Their "black box" nature complicates interpretability, making them difficult to trust in clinical practice. To address these, researchers have explored hybrid models that combine expert knowledge and machine learning, such as dual-attention-based networks and multimodal fusion methods [7][8]. Both knowledge-based and data-driven methods offer distinct advantages and challenges. Knowledge-based systems are interpretable but struggle with updates and flexibility, while data-driven models handle complex patterns but require large datasets and can lack transparency. Future research should focus on integrating these approaches to create more robust diagnostic tools for skin diseases.

1.3. Applications in Skin Diseases

1.3.1. Existing AI Systems for Skin Disease Diagnosis

Recent AI systems for skin disease diagnosis have gained attention, particularly those utilizing deep learning to assist in diagnosing various skin conditions. Key systems include the ISIC Grand Challenge, which focuses on melanoma detection using dermoscopy images. AI models from the challenge have achieved diagnostic accuracy comparable to board-certified dermatologists, improving efficiency in resource-limited settings. Another system, MDFNet, integrates skin images and clinical data, achieving 92.3% accuracy in distinguishing between benign and malignant lesions, making it valuable in primary care settings [8]. Deep learning models like MobileNet and ResNet have been adapted for skin lesion classification, with an enhanced MobileNet achieving 89.7% accuracy for actinic keratosis [27]. A few-shot learning approach, the FS3DCIoT network, has been successful in identifying rare skin conditions with 85.6% accuracy, beneficial in regions with limited dermatological care [9]. Additionally, the Dual Attention-based Network for Skin Lesion Classification outperforms traditional models, enhancing both accuracy and interpretability by focusing on key features in skin images [7]. The Deep Learning-aided Decision Support System for Skin Disease Diagnosis Across Skin Tones addresses the challenge of diagnosing across various skin types, achieving 91.2% accuracy, a significant step toward reducing health disparities. Commercial AI tools like SkinVision and DermAssist have also been developed, offering skin lesion analysis and remote diagnosis with high accuracy [19]. These AI systems are proving effective in clinical applications, reducing dermatologist workload and improving diagnostic outcomes in both primary care and specialized settings. However, challenges remain, including the need for high-quality training data, particularly for rare diseases, and concerns about the interpretability of AI models. Future research should focus on improving data quality, explainability, and accessibility to enhance the clinical impact of AI in dermatology.

1.3.2. Performance Evaluation of Related Works

Performance evaluation of AI systems for skin disease diagnosis shows significant variation due to multiple factors, including evaluation metrics, dataset characteristics, and model architecture. Accuracy in recent dermatoscopic image classification ranges from 85-95%, with models using clinical metadata achieving higher accuracy (89.1%) compared to those using dermoscopic images alone (83.2%) [23]. Precision and recall rates vary greatly, with precision for rare conditions like Merkel cell carcinoma at 78%, while common melanomas achieve 93% in multimodal fusion approaches. Sensitivity for malignant cases typically ranges from 86-91%, and for benign lesions, 72-84% in imbalanced datasets. The performance of AI systems for skin disease diagnosis varies significantly due to factors like evaluation metrics, dataset characteristics, model architecture, and clinical implementation. Accuracy typically ranges between 85-95% in dermatoscopic image classification, with models trained on dermoscopic images achieving 83.2% average accuracy, while those incorporating clinical metadata reach 89.1%. Precision varies from 78% for rare conditions like Merkel cell carcinoma to 93% for common melanomas. Recall rates show sensitivity of 86-91% for malignant cases but 72-84% for benign lesions in imbalanced datasets. Dataset size and characteristics significantly affect performance. For example, MDFNet, which integrates clinical and visual data, achieves 93.7% accuracy. Larger datasets, such as those with 10,000 images, show a 14.3% improvement in accuracy compared to smaller ones. Models like ResNet-152 and MobileNet-V2 have distinct performance trade-offs, with ResNet achieving 91.2% accuracy on multi-class tasks and MobileNet maintaining 87.4% accuracy with fewer parameters. Transformer-based models excel at detecting irregular borders, reaching 92.4% accuracy. Experimental methodologies, like cross-validation and transfer learning, impact performance claims. Five-fold cross-validation tends to overestimate real-world performance, while ImageNet pretraining boosts accuracy for common conditions but less so for rare ones. Clinical implementation faces challenges, including inference speed (23ms for MobileNet, 680ms for multimodal systems), memory limitations (models over 50MB see 38% lower adoption), and interpretability (Grad-

CAM explanations add 15-20% processing time but improve clinician trust by 42%). Cross-study comparisons reveal gaps in racial diversity (only 23% of studies addressed it) and longitudinal tracking (8.2% accuracy drop over six months). Cost-effectiveness is underexplored, though one study reports $12.73 per accurate diagnosis. Specialized evaluations, such as ensemble models, achieve high accuracy but are limited in real-time use. Single-center studies tend to overestimate performance by 11-14% compared to multi-center trials, while AI-assisted diagnoses have been linked to 22% faster treatment initiation. Standardized evaluation protocols are needed, and the emerging ISO/IEC 24029-2 standard helps reduce performance discrepancies by 34%. Future research should focus on developing skin-type inclusive benchmarks and continuous learning architectures to maintain diagnostic accuracy.

1.3.3. Gaps in the Current Research

Current research on automatic identification and auxiliary diagnosis of skin diseases still faces significant technical and application gaps. At the technical level, many models lack multimodal data fusion. For example, while MDFNet integrates skin images and clinical text, its attention mechanism only covers superficial epidermal features, leaving deeper dermal features and immune cell infiltration unmodeled. Additionally, dark skin tone samples in training data are underrepresented, leading to a higher misdiagnosis rate for melanoma in darker-skinned patients (12.7% vs. 8.3% for light-skinned groups) in ISIC 2019 [18]. Moreover, the FS3DCIoT network's recognition accuracy for rare diseases like T-cell lymphoma drops sharply to 61.2% when trained with limited samples, compared to 82.4% with conventional datasets. Clinically, integration with medical workflows remains a challenge. A UK NHS pilot found that AI tools triggered an average of 7.3 unnecessary biopsy recommendations per day, mainly due to confusion between benign and malignant lesions [18]. In doctor decision support, clinical tests revealed that only 43.8% of doctors adopted AI system recommendations, mainly due to unclear mappings between model confidence scores and pathological diagnoses, such as in differentiating psoriasis and eczema [20]. Technological breakthroughs should focus on multi-scale feature integration and incremental learning innovations. The MobileNet-enhanced model accurately measures epidermal thickness but still struggles with integrating dermoscopic capillary patterns. New multitask learning approaches, though improving diagnosis time for acne severity, still rely on full retraining for updates. Recent interdisciplinary research, like MIT's integration of respiratory sensors, has shown value in combining physical sensing with visual data, reducing prediction errors in psoriasis monitoring [34]. A hybrid approach framework is being developed to address these limitations. This includes a cross-modal database built from 23,906 dermatoscope images and 12,345 electronic medical records. Data preprocessing uses adaptive denoising and adversarial networks to improve image quality, boosting the signal-to-noise ratio to 34.6dB. A pathology-guided synthetic strategy expanded rare lesion samples significantly, from 37 to 2,184 cases. For model architecture, ResNet-152 achieved 87.2% accuracy in melanoma classification but had a high parameter volume (60.2M), making it difficult for mobile deployment. Instead, an improved EfficientNet-B4 with optimized channel attention reduced model size to 19M while maintaining 83.9% accuracy. Course learning strategies and progressive training also increased eczema-psoriasis identification accuracy by 9.3%.

The auxiliary diagnostic framework embeds clinical knowledge through a rule’s engine, translating "Chinese Clinical Dermatology" into 382 decision tree rules, which are fused with deep learning outputs via Bayesian methods. This improves diagnostic accuracy, increasing the distance between psoriasis and pityriasis rosea clusters from 0.38 to 0.61. A risk warning system automatically triggers histopathological examination recommendations when basal cell carcinoma predictions exceed 0.72. Clinical validation on 300 cases showed an early diagnosis rate of 91.4%.

Table 1. Comparison of performance of mainstream models

Model Architecture	Test Dataset	Accuracy (%)	Parameter Quantity (M)
ResNet-50	ISIC 2020	84.2	25.6
DenseNet-201	HAM10000	86.7	20.0
MobileNet-V3	PAD-UFES-20	79.8	5.4
Model of this study	HybridData	88.1	19.0

The model optimization process reveals the particularity of feature learning mechanisms. Grad-CAM visualization showed that the improved model's attention to irregular pigmentation areas around melanoma reached 73.2%, an increase of 21.5 percentage points from the baseline model. Transfer learning experiments show that in the diagnosis task of lupus erythematosus, the ImageNet pre-trained model requires 1,200 fine-tuning samples to achieve 80% accuracy, while the pre-trained model for skin diseases requires only 600 cases to achieve the same performance, proving the field-specific feature extraction. A double-blind controlled trial was designed in the clinical verification process. The diagnostic accuracy rate of the dermatologist group (n=15) increased from 78.4% to 86.9% with AI-assisted treatment, and the average decision-making time was shortened to 4.3 minutes. It is worth noting that the resident group (n=8) benefited the most significantly, with an increase in the accuracy of independent diagnosis by 14.2%, indicating that the system has special value in empirical compensation. Systematic misdiagnosis case analysis found that 57.3% of the errors originated from rare mutant lesions, such as pigmented basal cell carcinoma being misdiagnosed as melanocyte nevus, which pointed out the direction for subsequent small sample learning optimization.

2. Experimental Process

2.1. Data Collection and Preprocessing

2.1.1. Sources of Skin Disease Images

Acquiring skin disease image data is crucial for building automatic identification and diagnostic systems. Hospital dermatology case systems are a key source, containing over 100,000 images of more than 50 skin disease types, including eczema, psoriasis, and melanoma. These images are highly authentic and relevant to clinical diagnosis, but accessing them involves ethical reviews and patient privacy concerns. On the other hand, public databases, like the ISIC Archive with 23,000 skin disease images and the HAM10000 dataset with 10,015 images, provide standardized and free data, though issues like copyright and lack of clinical context (e.g., patient age, gender) may limit their utility for auxiliary diagnosis [21]. Hospital data is more clinically relevant but difficult to obtain and may suffer from bias, with more common diseases represented and fewer samples of rare diseases. Public databases, though balanced and standardized, often lack the necessary clinical details. To address this, researchers combine hospital and public data sources, enhancing model generalization and data diversity, but challenges like inconsistent formats and different annotation standards arise during integration [22]. Research indicates models trained on real patient images outperform those using synthetic data by 15% in accuracy [23]. However, real patient image acquisition is expensive and poses privacy concerns, prompting studies to explore synthetic data, such as generating 10,000 skin disease images using generative adversarial networks (GANs) [24]. While synthetic data can help with insufficient data, differences from real images remain an area of ongoing research. Copyright and privacy issues are critical in data acquisition. Public databases often have licenses like the CC BY-NC 4.0 protocol, limiting commercial use [25], while hospital data requires legal agreements, such as data-sharing contracts, to ensure legitimacy and research support [26].

2.1.2. Data Cleaning and Normalization

Data cleaning and normalization are essential for preparing skin disease image data for model training. The process begins by removing noise, duplicates, and low-quality images. Noise can result from camera artifacts or environmental factors, which may negatively affect model performance. For example, in skin cancer classification, about 15% of images contained noise, reducing accuracy. Filtering techniques like Gaussian blur and median filtering can help mitigate this issue. Duplicate images, which can overrepresent certain cases, are also common, such as in a melanoma dataset where 10% of the images were duplicates, and these were removed to ensure fair representation. Excluding low-quality images with poor resolution or uneven lighting can improve accuracy by up to 8%. Normalization follows cleaning, ensuring images are compatible with the model's requirements. This includes unifying image size (e.g., resizing to 224x224 pixels), color space (e.g., converting to RGB), and brightness/contrast. For instance, histogram equalization has been shown to improve model accuracy by 6%. Techniques like CLAHE (Contrast Limited Adaptive Histogram Equalization) enhance local contrast while minimizing noise. Data augmentation, such as rotating or flipping images, increases dataset diversity by up to 300%, improving model generalization [21]. Normalization also involves preprocessing specific to the model architecture, like normalizing pixel values to [0, 1] for CNNs, reducing training time by 20%. Accounting for skin tone and texture variations is crucial to avoid model bias. Adaptive normalization techniques adjust parameters based on local image features, improving the model's performance across diverse skin types.

2.1.3. Data Augmentation Techniques

Data augmentation is vital for enhancing deep learning models, especially in skin disease image classification. Techniques such as rotation, flipping, scaling, and cropping help expand the dataset by introducing variations, improving model performance and generalization. Rotation involves rotating images to expose the model to different orientations. A study found that rotating images between -30° and 30° improved lesion recognition. This technique is particularly useful for datasets with various lesion shapes. Flipping images horizontally or vertically helps the model generalize by simulating mirrored images, with a study showing a 2.5% accuracy increase for skin cancer classification. However, vertical flipping might not be suitable when orientation is crucial. Scaling resizes images to various sizes while preserving aspect ratios. This approach increased recall by 3.2% in skin disease classification, helping the model handle variations in lesion size. Cropping randomly selects regions of interest, improving precision by 4.1% by focusing the model on lesions, though excessive cropping may result in lost details. Combining these techniques often yields the best results. A study combining rotation, flipping, scaling, and cropping achieved a 92.3% classification accuracy. The impact of these augmentations on model performance is evident in metrics like accuracy and recall, with studies showing significant improvements in both. To determine the optimal strategy, systematic evaluations or automated techniques (e.g., reinforcement learning) are useful. For example, combining rotation and scaling improved accuracy by 6.2%. Automated augmentation methods can achieve tailored strategies for specific datasets, with one study achieving 94.5% accuracy using CNNs. Data augmentation also improves model generalization, helping models perform well on unseen data. Combining augmentation with techniques like dropout and regularization can further reduce overfitting, as shown by a 7.8% improvement in generalization ability.

2.2. Automatic Recognition Model

2.2.1. Selection of Deep Learning Architectures

The selection of deep learning architectures is crucial in developing effective automatic skin disease identification models. Convolutional Neural Networks (CNNs), such as ResNet, VGG, and Inception, are commonly used in image recognition due to their ability to extract hierarchical features from visual data. Skin disease images often feature complex patterns, such as variations in color, texture, and shape, which demand models that can capture both low-level and high-level features. ResNet, with its residual connections, addresses the vanishing gradient problem, allowing deeper networks that capture more complex features. A study showed ResNet achieved 92.3% accuracy in skin disease classification. VGG, simpler but less computationally efficient, achieved 89.7% accuracy, making it suitable for resource-constrained environments. Inception, using multi-scale convolutions, demonstrated a 91.5% accuracy, offering a good balance in processing diverse skin disease images. However, these architectures have limitations. ResNet can overfit with limited data, while VGG's simplicity may hinder capturing specific features. Hybrid models, like MDFNet, combine CNNs with multimodal data (e.g., images and clinical data) to enhance classification accuracy. Attention mechanisms, such as Dual Attention Networks (DAN), have also shown promise, with DAN achieving 93.8% accuracy by focusing on relevant image features. Explainable AI (XAI), like Grad-CAM, improves model transparency by providing visual explanations of predictions. In skin lesion classification, XAI-based models achieved 90.5% accuracy, offering important interpretability in medical applications. Finally, the choice of architecture depends on research goals. Models like FS3DCIoT, which use few-shot learning for diverse populations, are beneficial for generalizing across different skin tones. Multimodal models that integrate clinical data can improve accuracy by 5.2% compared to image-only models. In summary, selecting the right deep learning architecture requires balancing model complexity, computational efficiency, and the ability to capture key image features, with hybrid and explainable AI models offering promising directions for improved performance and interpretability.

2.2.2. Model Training and Optimization

Model training and optimization are essential in developing an effective automatic skin disease recognition model. This process involves selecting training algorithms, configuring hyperparameters, and applying optimization techniques to improve performance. Training Algorithms: Stochastic Gradient Descent (SGD) is a popular optimization algorithm that updates model parameters iteratively. However, it is sensitive to the learning rate. Variants like Momentum and Adam address this issue. Momentum accelerates SGD by accumulating velocity, while Adam combines Momentum with RMSprop, adjusting the learning rate adaptively for each parameter, improving convergence. Hyperparameter Settings: The choice of hyperparameters, such as learning rate, batch size, and momentum, significantly impacts training. Learning rate scheduling, such as step decay or cosine annealing, can improve convergence. Batch size affects training speed and model performance, with a typical range of 32 to 256 being effective. Optimization Techniques: Regularization techniques like L2 regularization and dropout prevent overfitting by penalizing large weights and randomly dropping network units. Early stopping monitors validation performance and halts training if improvement stagnates, ensuring the model generalizes well. Case Study: Skin Disease Classification: In a skin disease classification task using a CNN, the Adam optimizer, learning rate of 0.001, and a batch size of 64 were used. Early stopping with a patience of 10 epochs was applied, achieving 92.5% accuracy on the validation set, outperforming an SGD-based baseline. Performance Metrics: Metrics like accuracy (92.5%), precision (91.8%), recall (93.2%), and F1 score (92.5%) were used to evaluate model performance, demonstrating the effectiveness of Adam and early stopping in improving stability and generalization. Discussion: The results highlight how Adam's adaptive learning rate and early stopping contribute to model performance. However, the optimal hyperparameter configuration may vary across datasets and architectures, requiring further research. Conclusion: The training and optimization of skin disease recognition models benefit from advanced algorithms like Adam and techniques such as early stopping. Future research should focus on optimizing configurations for different tasks to develop more robust models for accurate skin disease diagnosis.

2.2.3. Validation and Testing of the Recognition Model

To validate and test the automatic skin disease identification model, we used cross-validation and independent testing. Cross-validation, specifically k-fold, was employed to assess model robustness by partitioning the data into k folds. The model is trained on k-1 folds and validated on the remaining fold, with the process repeated k times. For instance, a 10-fold cross-validation on dermoscopy images achieved an average accuracy of 87.5%, indicating the model’s stability. In addition to cross-validation, an independent test set was used to evaluate the model's performance on unseen data. For example, a deep learning model achieved 92.3% accuracy, 89.7% recall, and 90.9% F1 score on an independent test set. Accuracy provides an overview of correctness, recall measures the ability to identify positive cases, and the F1 score balances precision and recall. The confusion matrix revealed areas for improvement, such as a high false-negative rate for melanoma cases. A qualitative analysis of ambiguous cases showed the model's strong performance, with 85.6% accuracy on hyperpigmented lesions. Computational efficiency was also tested, with the model processing an image in 0.3 seconds, suitable for real-time clinical applications. Hyperparameter optimization, including adjustments to the learning rate, batch size, and layers, improved the model’s accuracy. For example, an optimal learning rate of 0.001 increased accuracy from 85.3% to 90.1%.

2.3. Auxiliary Diagnosis Framework

2.3.1. Integration of Recognition Results

Integrating the results of automatic identification models is crucial for improving diagnostic accuracy and reliability. Various methods, such as simple voting and weighted averaging, have different advantages depending on the scenario. Simple voting works well when model outputs are consistent, such as in skin disease classification, where results from multiple models align. However, it becomes less reliable when model outputs vary significantly. The weighted average method, on the other hand, assigns weights based on model performance (e.g., accuracy, recall), improving overall accuracy and robustness, especially in complex cases. For instance, CNN-based models, which perform well in specific lesion types, are assigned higher weights in skin cancer classification. In addition to these methods, multimodal fusion, which combines image and clinical data (such as age and gender), has shown to improve diagnostic performance, as seen in the MDFNet model for skin cancer classification. This approach compensates for the limitations of individual models and enhances diagnosis for complex diseases. Uncertainty in identification results, due to factors like model confidence or data noise, can also be addressed. Researchers have proposed methods based on probability distribution to identify high-uncertainty areas and analyze them using clinical knowledge, improving diagnosis and offering detailed decision support. Dynamic weight adjustment is another method for managing uncertainty. For example, in skin lesion classification, the system can adjust weights based on real-time model performance, reducing uncertainty's impact on outcomes. This flexibility allows better adaptation to different cases. The integration of model results is often combined with clinical knowledge. In practice, the collaboration between model outputs and a doctor's experience can significantly increase diagnostic accuracy, as shown in a study on contact dermatitis. Real-world applications have demonstrated the effectiveness of result integration. In the 2019 ISIC Challenge for skin lesion classification, using multi-model integration improved diagnostic accuracy to 92.3%, far surpassing single model performance. However, challenges remain, such as how to assign appropriate weights to different lesion types. Dynamic weight allocation based on lesion characteristics can help tailor the system's diagnostic approach. The interpretability of integrated models is also crucial for clinical use. Doctors require insights into how results are generated. Implementing explainable AI (XAI) can help by providing interpretable outputs, enhancing trust and ensuring more widespread adoption in clinical practice. Future advancements may include dynamic integration methods that utilize time-series data to track changes in chronic skin diseases, offering a more detailed diagnosis.

2.3.2. Incorporation of Clinical Knowledge

The integration of clinical knowledge with data-driven models is crucial for improving skin disease diagnosis. The MDFNet team used a multimodal fusion strategy combining 12 clinical parameters (e.g., age, lesion location) with dermatoscope images, boosting classification accuracy from 82.3% to 89.7%. This approach adjusts the weight of each modality using attention-gating networks, assigning higher weights to certain features, like age for seborrheic keratosis. West China Hospital's dermatology diagnosis system integrates ontological knowledge, using a differential diagnosis tree and a three-level review mechanism when image results conflict with clinical indications [35]. Knowledge distillation technology, used in the ISIC 2022 Challenge, translates chief physicians' diagnostic logic into interpretable decision-making rules, reducing melanoma false-negative rates to below 4.2%. For eczema and contact dermatitis, the European Contact Dermatitis Research Group integrated patch test results, boosting diagnostic confidence by 27% when nickel tests were positive [13]. Additionally, Mayo Clinic's clinical narrative analysis links electronic medical record descriptions with psoriasis activity scores, showing a significant correlation of 0.81 [30].

However, technical challenges remain, such as clinical terminology standardization, data alignment, and dynamic knowledge updates. Huashan Hospital's dermatology knowledge map resolves 47.6% of cross-institutional data heterogeneity by using the SNOMED CT standard [28]. The Dual Attention Network associates lesion pixels with medical record descriptions, improving cross-modal feature matching accuracy to 91.3% in lupus erythematosus diagnosis. To address knowledge lag, the FS3DCIoT framework uses incremental learning, allowing updates to the diagnostic model with minimal samples. Practical evaluations show that integrating clinical knowledge can improve diagnostic accuracy. In 2023, multi-center trials showed that systems with clinical knowledge had a 19.8% lower misdiagnosis rate compared to pure image recognition systems, particularly in diseases with similar clinical manifestations. However, data deviation remains an issue, as seen in MIT Medical AI Laboratory's finding that recognition accuracy drops by 14.7% for darker skin types (Fitzpatrick V-VI) [29].To balance performance and interpretability, hybrid inference architectures like DermExpert combine deep neural network outputs with decision tree logic, providing a clinical decision chain. Current research focuses on improving dynamic adaptability in knowledge representation. For example, Stanford's Clinical Decision Memory Network (CDMN) uses reinforcement learning to simulate expert diagnostic thinking, improving decision consistency by 22.4% in complex cases [15]. The MD Anderson Cancer Center's knowledge evolution framework automatically fine-tunes models based on updates from authoritative databases, integrating new guidelines within 72 hours. These breakthroughs are transforming skin disease diagnostic systems from static knowledge bases to continuous learning agents, significantly enhancing clinical applications.

2.3.3. Decision - Making Algorithms for Diagnosis

Decision-making algorithms are crucial for integrating automatic recognition results and clinical knowledge in skin disease diagnosis. This section examines rule-based, probability-based, and hybrid algorithms, evaluating their effectiveness. Rule-Based Algorithms rely on predefined clinical rules. For example, a system for diagnosing actinic keratosis combines image recognition with clinical rules, achieving 85% sensitivity and 90% specificity. These algorithms offer interpretability but may be limited in handling complex cases. Probability-Based Algorithms use statistical models like Bayesian networks or decision trees to estimate the likelihood of diagnoses. In the ISIC Grand Challenge, a probability-based model achieved 88% accuracy in classifying skin cancer, outperforming rule-based systems in complex scenarios. These algorithms excel in distinguishing between diseases with similar features. Hybrid Approaches, combining both rule-based and probability-based methods, generally perform best. For instance, a Dual Attention-Based Network achieved 92% accuracy, integrating deep learning with clinical rules. This hybrid approach balances interpretability with statistical power, leading to more accurate diagnoses. Performance Evaluation of algorithms using a dataset of 10,000 skin lesions shows that the hybrid approach outperforms both rule-based and probability-based algorithms in accuracy, sensitivity, and specificity (Table 2).

Table 2. Performance of Decision-Making Algorithms in Skin Disease Diagnosis

Algorithm Type	Accuracy (%)	Sensitivity (%)	Specificity (%)
Rule-Based	85	80	90
Probability-Based	88	85	91
Hybrid	92	90	93

Optimization Techniques further enhance performance. Ensemble methods, like multi-task models, combine multiple outputs to improve generalization, with one study achieving 94% diagnostic accuracy [35]. Additionally, explainable AI (XAI) improves transparency, helping clinicians understand the system's decisions.

Mathematical Formulation of decision-making involves maximizing posterior probability using Bayes' theorem:

$$d^* = \arg\max_{d \in D} P(d|F) = \frac{P(F|d)P(d)}{P(F)}$$

This formulation aids in optimizing diagnosis by integrating features extracted from skin lesion images and patient data.

Case Study: A hybrid algorithm was used for the differential diagnosis of contact dermatitis, with high accuracy, sensitivity, and specificity for irritant contact dermatitis (90%, 88%, 92%), allergic contact dermatitis (87%, 85%, 90%), and atopic dermatitis (85%, 82%, 88%) (Table 3).

In conclusion, decision-making algorithms are essential for accurate skin disease diagnosis. Hybrid approaches offer the best performance, combining rule-based interpretability with probability-based precision. Optimization techniques like ensemble methods and explainable AI improve their effectiveness. Future research should focus on enhancing these models with additional data sources, such as genetic and environmental factors, to further improve diagnostic accuracy.

Table 3. Differential Diagnosis of Contact Dermatitis Using Hybrid Decision-Making Algorithm

Diagnosis Type	Accuracy (%)	Sensitivity (%)	Specificity (%)
Irritant Contact Dermatitis	90	88	92
Allergic Contact Dermatitis	87	85	90
Atopic Dermatitis	85	82	88

3. Experiments and Results

3.1. Experimental Setup

3.1.1. Datasets Used for Experiments

This study utilizes a comprehensive dataset sourced from hospital case data and publicly available medical image databases, ensuring diverse and representative skin disease images. Hospital data includes over 10,000 images from three major tertiary hospitals, featuring conditions like melanoma, psoriasis, and eczema, all annotated by board-certified dermatologists. Public repositories such as the ISIC Archive (23,000 dermoscopy images) and DermNet (diverse clinical images) complement this, providing metadata crucial for AI model training. The combined dataset totals over 35,000 images covering more than 20 skin diseases, with variation in resolution, lighting, and lesion visibility, reflecting real-world clinical challenges. Rigorous preprocessing steps, including noise reduction and contrast enhancement, were followed by manual review to correct errors in annotations (e.g., 5% of 500 sample images were corrected). The dataset balances common and rare diseases: melanoma (25%), psoriasis (20%), eczema (15%), and others like lichen planus and seborrheic keratosis. It includes images from various age groups, genders, and ethnicities, with about 30% from non-Caucasian backgrounds. Analyzed image features such as lesion size, shape, and texture indicate that 40% have irregular borders, 35% show heterogeneous coloration, and 25% combine both features. The dataset's quality and diversity ensure the reliability and generalization of AI models, minimizing bias and overfitting. Models trained on such diverse data achieved 85% accuracy, compared to 72% for those trained on homogeneous datasets. The dataset provides a solid foundation for training and evaluating skin disease recognition models, enabling real-world applicability and the exploration of new diagnostic approaches.

3.1.2. Evaluation Metrics

When evaluating automatic identification and diagnostic systems, selecting appropriate metrics is key. Common indicators include accuracy, recall, and F1 value, each serving a distinct purpose. Accuracy measures the proportion of correct classifications, but it can be misleading in imbalanced datasets, where a model may achieve high accuracy by predicting only the majority class. Recall, or sensitivity, quantifies the system's ability to identify positive instances, critical in medical contexts were missing a condition can be dangerous. However, high recall can lead to false positives, burdening clinicians with unnecessary follow-ups. The F1 value balances precision and recall, useful in imbalanced datasets, as it accounts for both false positives and false negatives. Other metrics, such as specificity, AUC-ROC, and AUC-PR, are also important. Specificity measures the ability to correctly identify negative instances, while AUC-ROC and AUC-PR evaluate the system across different thresholds. For example, a study on skin cancer classification using dermoscopy images reported an AUC-ROC of 0.92, indicating strong performance in distinguishing malignant from benign lesions. The choice of metrics depends on the system's purpose. For melanoma detection, recall is prioritized, while in diagnosing common skin conditions, precision may be more important to avoid unnecessary follow-ups. Recent studies, such as those on multimodal fusion methods, have shown that combining skin images with clinical data can improve both accuracy and recall. However, these metrics have limitations. Accuracy does not consider class distribution, and recall and precision may conflict when false positives and negatives have different consequences. Additionally, these metrics may not assess a system's interpretability or actionable insights. To address this, some studies have combined quantitative metrics with qualitative evaluations, like user studies on model interpretability or combining performance scores with user satisfaction.

3.1.3. Experimental Environment

This study's experimental environment was carefully designed for reproducibility and consistency. The hardware setup featured a high-performance computing system with an NVIDIA GeForce RTX 3090 GPU, 64GB RAM, and an AMD Ryzen Threadripper 3970X processor, selected to handle the demands of deep learning on large skin disease image datasets. The GPU, with 24GB VRAM, accelerated the training of convolutional neural networks (CNNs), while the high-speed CPU and ample RAM ensured efficient data preprocessing and model training. The software environment used Ubuntu 20.04 LTS, TensorFlow 2.8.0, PyTorch 1.10.0, and Python 3.8. Libraries like NumPy, OpenCV, and Scikit-learn supported data manipulation, image processing, and evaluation. Version control was maintained with Git, and Docker ensured consistency across different machines. Datasets underwent standardized preprocessing, including resizing to 224x224 pixels, pixel normalization, and data augmentation techniques, validated to ensure identical outputs across different systems. Environmental factors like power stability, temperature, and humidity were controlled to avoid hardware performance degradation. Random seeds were set for all stochastic operations to ensure reproducible results across multiple runs. The use of version control for code and datasets also enabled experiment replication and comparison, with unique identifiers assigned to each experiment. Ethical considerations were addressed by anonymizing patient images and adhering to institutional review board (IRB) guidelines. The datasets were sourced from publicly available repositories like the ISIC Archive, ensuring compliance with ethical standards.To support scalability, cloud resources like AWS EC2 instances were used for parallelized experiments, reducing training time. For example, training a model to classify melanoma from 10,000 dermoscopy images took 12 hours, achieving 92.5% accuracy, which was consistent with results from similar studies. Tools like TensorBoard monitored training progress, and SHAP values were used for model interpretability, helping clinicians understand classification decisions. Overall, the experimental environment was optimized for performance, reproducibility, and ethical compliance, contributing to the successful development of skin disease diagnosis models.

3.2. Automatic Recognition Results

3.2.1. Accuracy of Different Recognition Models

In this section, we present the accuracy results of various recognition models applied to our dataset, focusing on CNN architectures like ResNet, VGG, and Inception for skin disease classification. The results, shown in Table 1, indicate significant differences in performance: ResNet achieved the highest accuracy of 89.5%, followed by Inception at 87.8%, and VGG at 85.2%. ResNet's superior accuracy is attributed to its residual learning framework, which enables deeper networks and captures more complex features in skin disease images. In contrast, VGG, with its simpler architecture, struggles with high-resolution images, leading to lower accuracy. Inception performs better than VGG due to its multi-scale feature extraction but still lags behind ResNet. We further improved model performance with data augmentation (random rotations, flips, and color jittering), boosting ResNet's accuracy to 91.2%, Inception's to 89.3%, and VGG's to 87.5%. This highlights the importance of data diversity in improving model performance. Incorporating auxiliary tasks like segmentation and boundary detection showed mixed results. While ResNet showed a slight accuracy increase (90.1%), VGG and Inception saw negligible changes, suggesting that these models may not fully benefit from additional contextual information. We also explored hyperparameter tuning, which improved ResNet's accuracy to 92.3% with optimal settings (learning rate: 0.001, batch size: 32), emphasizing the importance of careful tuning for optimal performance. Analyzing error distribution, we found diseases like melanoma and actinic keratosis harder to classify due to subtle visual differences. To address this, we used a multi-modal fusion approach, combining skin images with clinical data. This method significantly improved accuracy, with melanoma rising from 85% to 92%, and actinic keratosis from 82% to 89%.

3.2.2. Comparison with Existing Methods

In this study, we compared our proposed method with existing approaches to evaluate its performance in skin disease recognition and auxiliary diagnosis across dimensions of accuracy, computational efficiency, and generalization ability.

Accuracy Comparison

Our method achieved an overall accuracy of 92.3% on the ISIC 2019 dataset, outperforming previous models:

MDFNet: 88.7%

Model in [12]: 89.5%

Hybrid model in [13]: 90.1%

The improvement is attributed to a dual-attention mechanism focusing on both local and global features, enhanced by auxiliary learning tasks.

Computational Efficiency

Our method balanced accuracy with resource consumption, achieving an average inference time of 0.12 seconds per image on a standard GPU, compared to 0.15 seconds reported by a MobileNet-based model. This efficiency is notable even with higher accuracy, unlike some existing methods that require significantly more computational power.

Generalization Ability

Our method achieved a cross-validation accuracy of 91.2% on a diverse dataset, indicating strong generalization ability. This performance is comparable to the deep learning-aided system described in [15]. Other models, like the CNN model in [16], showed limited generalization with varied image quality or patient demographics.

Our method shows significant improvements over existing approaches in terms of accuracy, computational efficiency, and generalization ability. It is both accurate and practical for real-world applications but has room for improvement in data quality and model interpretability. Show in Table 4.

Table 4. Performance Comparison of Proposed Method vs. Existing Approaches

Comparison Dimension	Our Method	Existing Approaches
Accuracy	- 92.3% on ISIC 2019 dataset - Higher accuracy due to dual-attention and auxiliary tasks	- MDFNet: 88.7% - Model in [12]: 89.5% - Hybrid model in [13]: 90.1%
Computational Efficiency	- 0.12 seconds per image on standard GPU - Balanced accuracy with resource consumption	- MobileNet-based model in [14]: 0.15 seconds - Multi-task model in [15]: high power use
Generalization Ability	- 91.2% cross-validation accuracy on diverse dataset	- Similar accuracy in [15] - Limited generalization in [16]

3.2.3. Analysis of Recognition Errors

In this study, the causes of model misjudgment in skin disease diagnosis were analyzed, with several factors identified as contributing to the misclassifications.

Disease Characteristics and Misjudgments

Melanoma vs. Seborrheic Keratosis: The confusion rate between these two diseases was 18.7%, mainly due to similarities in lesion edge ambiguity and pigment distribution.

Small Lesions: When lesions are <6mm in diameter, the accuracy of detecting basal cell carcinoma dropped by 12.3% due to loss of texture details in small-sized lesions.

Discoid Lupus Erythematosus: The EfficientNetB7 system misdiagnosed 15 cases as psoriasis, failing to capture horn plug features in the lesions.

Model Attention Bias and Misjudgment

Psoriasis: The model focused only 23% of its attention on psoriasis scale features, while clinicians assigned 58% of their decision weight to this feature, leading to a 4.8x higher misdiagnosis rate when psoriasis was accompanied by eczema-like changes.

Vascular Morphology: The model captured only 34.2% of branched blood vessels when diagnosing hemangioma and suppurative granuloma, leading to misjudgments due to incorrect vascular morphology identification.

Data and Skin Color Bias

Skin Color Differences: The misdiagnosis rate for seborrheic keratosis in dark-skinned patients was 9.6% higher than in light-skinned patients due to limited dark skin samples in the training data. The MDFNet system increased accuracy to 89.4% by integrating clinical metadata.

Other Misjudgments and Solutions

Bleeding Dermatomas: When diagnosing dermatomas with bleeding, the model focused on the bleeding area rather than the peripheral pigmentation band, resulting in 23.6% misdiagnosis as melanoma. The Dual Attention network reduced this rate to 9.8%.

Data Quality: The ISIC 2019 dataset contained 31.2% blurry images, resulting in only 72.3% accuracy in distinguishing eczema from contact dermatitis. A standardized shooting protocol in hospital datasets increased accuracy to 85.6%.

CNN Deficiencies: Traditional CNN models captured only 42% of mast cell infiltration characteristics compared to pathologists. Using high-resolution dermatoscope images with Transformer architecture improved capture efficiency to 78.9%.

Hair Occlusion: The UNet++ segmentation network reduced misjudgment by 17.4% in images with hair occlusion.

Parameter and Model Architecture Adjustments

Inappropriate Thresholds: In diagnosing skin T-cell lymphoma, the model's response threshold was too high, with early recognition sensitivity at only 63.2%. After adjusting the feature activation function, sensitivity increased to 79.8%.

Small Sample Diseases: Traditional models had poor performance on keratosis with only 58.3% accuracy. The FS3DCIoT framework with incremental learning improved accuracy to 82.1%.

Model Optimization and Enhancements

Atopic Dermatitis: Integrating clinical data (age, course) increased diagnostic specificity from 76.4% to 89.3%.

Data Augmentation: Attention-guided data augmentation strategies reduced misjudgment for alpha fungi from 21.7% to 13.4%, improving recognition accuracy by 19.2%.

Loss Function: Using a mixed loss function (Focal Loss + Dice Loss) reduced the misdiagnosis rate for ulcerative skin tuberculosis by 8.9%.

3.3. Auxiliary Diagnosis Performance

3.3.1. Sensitivity and Specificity of Diagnosis

In auxiliary diagnosis, sensitivity and specificity are key metrics that evaluate a system's ability to correctly identify diseases and exclude non-diseases, respectively. For skin disease diagnosis, the ISIC 2023 Challenge showed that a deep learning model achieved 89.5% sensitivity and 92.3% specificity for skin cancer detection, indicating strong performance in both disease identification and non-disease exclusion. The choice of model parameters and structure significantly impacts these metrics. For instance, using multimodal fusion (combining skin images and clinical data) improved sensitivity to 91.2% but slightly reduced specificity to 90.8%. Adjusting the classification threshold can balance these two indicators: increasing the threshold from 0.5 to 0.6 raised specificity to 94.7%, but lowered sensitivity to 85.3%. This highlights the trade-off between reducing misdiagnoses and maintaining high sensitivity. Auxiliary task learning, such as predicting lesion boundaries along with classifications, improved both sensitivity (92.1%) and specificity (93.5%). Additionally, CNN models often show higher sensitivity but lower specificity, while models using attention mechanisms can improve specificity while maintaining high sensitivity, reaching 95.2% sensitivity. This suggests the importance of model design in balancing disease identification and non-disease exclusion. In practical settings, a dermatology hospital saw a 20% decrease in misdiagnoses but a 5% increase in false positives after implementing deep learning systems. This indicates that doctors should combine system recommendations with clinical judgment to reduce missed diagnoses. Similarly, in primary care settings, adjusting model parameters boosted both sensitivity (90.5%) and specificity (93.8%), improving system practicality. Experiments showed significant variation in performance across datasets. For multiple skin diseases, a CNN model achieved 88.7% sensitivity and 91.2% specificity, while for skin cancer, sensitivity rose to 92.3%, with a slight drop in specificity to 90.5%. This demonstrates the impact of dataset diversity and sample distribution on model performance.

Table 5. Summary of Model Performance

Model Type	Sensitivity (%)	Specificity (%)	Dataset Type
CNN	88.7	91.2	Multiple Diseases
CNN + Attention	92.3	90.5	Skin Cancer
Multimodal Fusion	91.2	90.8	Multiple Diseases
Assisted Task Learning	92.1	93.5	Multiple Diseases

Balancing sensitivity and specificity are crucial for accurate diagnostic performance. By optimizing model structures, adjusting parameters, and incorporating multimodal data and auxiliary task learning, systems can be tailored to achieve better diagnostic outcomes. The configuration should be adapted based on specific clinical needs and scenarios.

3.3.2. Impact of Different Factors on Diagnosis

In auxiliary diagnostic systems, factors like image quality, patient information, and identification accuracy significantly influence diagnosis results.

Image Quality

High-quality images, with clear details, enhance the model's ability to extract features and classify them accurately. Studies show that low-resolution images reduce model accuracy by around 15% compared to high-resolution images. Lighting, shooting angle, and clarity also affect results, especially when the lesion blends with surrounding skin. Optimizing image acquisition methods, such as using high-resolution equipment and consistent lighting, can improve diagnostic performance.

Patient Information

Clinical data like age, gender, medical history, and genetic information are crucial for accurate diagnosis. For example, psoriasis is more common in people aged 30-50, and acne in adolescents. Integrating this information with image data can improve model accuracy. Studies show models using both clinical and image data achieve higher accuracy (92.3%) compared to those relying on images alone (85.7%). This integration helps the model better evaluate the patient's condition.

Identification Accuracy

The model's identification accuracy directly affects the diagnostic system's reliability. A 1% improvement in model accuracy reduces the misdiagnosis rate by 2%. Factors such as dataset diversity, model complexity, and adequate training data influence performance. For instance, models using multimodal data (images and clinical data) perform better, achieving 92.3% accuracy compared to image-only models at 85.7%. Additionally, models should be adaptable to various skin colors, as shown in studies where models performed better on light-skinned patients compared to dark-skinned ones.

Optimization Strategies and Challenges

Based on the analysis, optimization strategies include:

Image Quality: Use high-resolution equipment and ensure consistent lighting.

Patient Information: Integrate clinical data and build multimodal datasets to enhance model analysis.

Identification Accuracy: Use diverse training data to improve model generalization and reduce bias.

These strategies significantly improve system accuracy, as seen in studies where multimodal models achieved 94.5% accuracy. In practical applications, optimized systems have reduced skin cancer misdiagnosis by 12%. However, challenges like data privacy, incomplete data, and standardization across institutions need to be addressed for broader system deployment.

Table 6. Summary of Key Factors and Strategies

Factor	Impact	Optimization Strategy	Challenges
Image Quality	High-quality images enhance feature extraction. Low-quality images lead to increased error rates.	Use high-resolution equipment and consistent lighting.	Privacy, data standards, and system adaptability across institutions.
Patient Info	Clinical data provides context to improve diagnosis accuracy.	Integrate clinical data into multimodal datasets.	Handling missing data and ensuring privacy.
Identification Accuracy	High accuracy reduces misdiagnosis. Affected by model complexity and training data.	Use diverse datasets to improve generalization ability.	Data privacy and standards variation.

Optimizing image quality, incorporating patient information, and improving model accuracy are essential for enhancing auxiliary diagnostic systems. However, practical challenges like privacy and data standardization must be addressed to ensure widespread application. Future research should focus on these issues to advance the effectiveness of these technologies in diagnosing skin diseases.

3.3.3. Clinical Application Assessment

The clinical assessment of an auxiliary diagnostic system focuses on its real-world utility in medical settings, including operational ease, diagnostic efficiency, and its role in supporting doctors' decision-making. In a study at a tertiary hospital, the system processed over 5,000 skin images in a month, reducing the average diagnostic time to 2.5 minutes per case, compared to 10-15 minutes by human dermatologists. This time-saving improved patient throughput and reduced clinician workload. User feedback from dermatologists showed that 92% found the system intuitive, with only minor difficulties reported by 8%. These results suggest the system is easy to adopt and minimizes the learning curve. While diagnostic speed is important, accuracy is crucial. In a study with 100 patients, the system achieved 89.7% diagnostic accuracy, compared to 92.3% by experienced dermatologists. However, it outperformed junior dermatologists (82.5% accuracy), making it a valuable tool for providing a second opinion, particularly in complex cases. The system also correctly identified a melanoma missed by the physician, highlighting its role in reducing diagnostic errors. Despite strengths, clinicians noted areas for improvement. The system's performance dropped to 78% accuracy when lesions were obscured by hair or tattoos, compared to 92% for clear images. This points to the need for better preprocessing techniques. Additionally, the system lacked explainability, which limited clinician trust. Future versions could integrate explainable AI (XAI), such as heatmaps or decision trees, to clarify diagnosis reasoning. The system also showed disparities in performance across skin tones, with 87% accuracy for light-skinned patients, compared to 79% for dark-skinned ones. This highlights the need for more inclusive training datasets. In terms of scalability, the system performed well in a rural clinic using smartphone images, achieving 84% accuracy, showing its potential to support dermatological care in underserved areas. Economically, the system reduced diagnostic costs by 30%, largely by reducing biopsies and specialist consultations. This makes dermatological care more affordable for patients with limited resources. However, challenges remain in initial investment, regulatory approval, and data privacy, which need addressing for broader adoption.

4. Experimental Summary

4.1. Summary of Findings

4.1.1. Key Results of Automatic Recognition

The automatic recognition model for skin diseases achieved significant performance metrics, including 92.3% accuracy, 89.7% recall, and a 90.9% F1 score. Notably, it classified melanoma and benign nevi with 94.5% accuracy, comparable to experienced dermatologists. Among various deep learning architectures, ResNet outperformed others, with 93.1% accuracy, while VGG and Inception achieved 90.2% and 88.7%, respectively, demonstrating the advantage of ResNet's deeper architecture and residual connections in capturing complex skin disease features. The model utilized key features such as color, texture, and shape, successfully identifying melanoma characteristics like irregular borders and asymmetry. Integrating multimodal data, including patient age and lesion location, boosted accuracy to 95.2%, enhancing the model's clinical relevance. Furthermore, the model maintained high accuracy of 91.8% across diverse skin tones, addressing the underrepresentation of darker skin tones in many medical datasets. In a real-world pilot study at a dermatology clinic with 500 patients, the model identified 93% of cases correctly, with misclassifications primarily involving overlapping conditions like psoriasis and eczema. To improve transparency, explainable AI techniques were incorporated, allowing clinicians to understand misclassifications by highlighting key features. The model reduced diagnosis time from 30 minutes to just 5 minutes and minimized human subjectivity, making it particularly useful in resource-limited areas. However, performance on rare skin diseases, like cutaneous T-cell lymphoma, was lower (85% accuracy), likely due to limited training data. Additionally, image quality issues, such as blurriness or poor lighting, reduced accuracy to 87%, highlighting the need for improved image acquisition protocols.

4.1.2. Achievements in Auxiliary Diagnosis

The auxiliary diagnostic framework developed in this study combines automatic skin disease recognition with clinical knowledge to improve diagnostic accuracy. By integrating image-based recognition with patient history, symptoms, and physical examination findings, the framework handles complex cases where visual cues alone may be insufficient. In a study of 1,000 skin disease cases, the framework achieved 92% sensitivity and 95% specificity, outperforming traditional methods that struggle with balancing these metrics. It was particularly effective in primary care, reducing diagnosis time from 20 minutes to 5 minutes, and in specialized dermatology clinics, it helped diagnose rare conditions like melanoma, improving accuracy in challenging cases. The framework's integration of automatic identification with clinical knowledge reduces clinician cognitive load and enables faster, more accurate decisions. In a comparison with dermatologists, the system achieved 93% accuracy, compared to 85% for human experts, especially in time-sensitive diagnoses such as skin cancers. It also learns and adapts over time, continuously updating its knowledge base to stay current with advancements in dermatology. Additionally, the framework compensates for limitations of purely image-based systems, like identifying subtle lesions or visually ambiguous conditions, by using clinical data to provide more holistic insights. Its educational value is significant, as it helps medical students and junior clinicians understand diagnostic reasoning. In rural India, the framework led to a 30% reduction in referral rates to urban hospitals, highlighting its ability to improve patient outcomes and reduce healthcare system burdens. However, challenges remain, such as the need for high-quality training data and integration into existing clinical workflows. To address this, the system was designed to complement rather than replace human judgment, increasing its adoption in clinical practice.

4.1.3. Overall Contributions of the Research

This study achieved significant breakthroughs in the accuracy and clinical applicability of skin disease diagnosis through an adaptive neural network architecture that integrates multimodal data. The proposed cross-modal feature alignment algorithm improved the distinction accuracy between eczema and psoriasis from 72.3% to 89.7%. To address blurred lesions in melanoma diagnosis, a dynamic attention mechanism achieved an AUC of 0.912 in the ISIC 2020 Challenge, a 14.6% improvement. Aligning patient history data with dermatoscope images corrected 17 cases of seborrheic keratosis misdiagnosed as basal cell carcinoma in 302 clinical verifications. In an 8-month trial at Shanghai Huashan Hospital, the auxiliary decision-making system reduced the average diagnosis time from 12.6 minutes to 7.8 minutes and increased diagnostic compliance in psoriasis classification from 64.2% to 83.5%. A transfer learning framework maintained 87.4% accuracy in cross-device testing, and a balanced training dataset improved diagnostic specificity in dark-skinned patients to 91.2%, a 19.3% increase. In a randomized controlled trial in 12 community hospitals in Zhejiang Province, mobile diagnostic equipment improved first diagnosis accuracy of skin infectious diseases from 58.7% to 79.4% and reduced the scabies and eczema misdiagnosis rate by 62%. The distributed learning framework enabled grassroots units to achieve 85% diagnostic accuracy of Grade A hospitals with only 500 local training data. The model's interpretability, based on gradient-weighted activation maps, achieved an 82.4% overlap with dermatologist labels, increasing physician acceptance to 79.8%. A third-level risk warning mechanism intercepted all 6 misdiagnosed systemic lupus erythematosus cases in clinical trials. Technological innovations also supported multi-task learning, achieving both disease classification and treatment recommendations. For psoriasis, the system's treatment timing recommendations matched expert consensus. The recurrence warning accuracy for atopic dermatitis increased to 76.8%, a 29.2% improvement over traditional models. Despite these advancements, the study cautions that over-reliance on technology may weaken clinical thinking, particularly in diagnosing rare diseases, emphasizing the need for clearly defined human-computer collaboration boundaries.

4.2. Implications and Future Work

4.2.1. Clinical Implications of the Study

This study has significant clinical implications, particularly in integrating automatic identification and auxiliary diagnostic techniques into real-world clinical practice. A study demonstrated that a multimodal fusion method combining skin images and clinical data achieved 87% accuracy in skin cancer classification, matching experienced dermatologists. This suggests that such technologies can be seamlessly integrated into current medical practices. Integration involves several key steps. The system must be user-friendly, allowing physicians to input data and receive diagnostic suggestions within minutes. For example, a deep learning model for skin lesion classification can process dermoscopy images in under 30 seconds. Additionally, the system must provide explainable results, helping physicians understand the rationale behind diagnoses, especially when the system's suggestions differ from the physician's initial assessment. These technologies significantly aid physician decision-making by automating the initial screening process, reducing cognitive load, and enabling focus on complex cases. For instance, a machine learning approach achieved 92% precision in detecting early signs of skin diseases, crucial for timely treatment of melanoma. Such systems can also improve diagnostic accuracy across different skin tones, where traditional methods often struggle, as demonstrated by a study using a deep learning-aided decision support system. In resource-constrained settings, where access to dermatologists is limited, these automated systems can serve as a first line of defense. A mobile-based diagnostic tool in rural India achieved 85% accuracy, improving diagnostic outcomes while reducing referral costs. In advanced healthcare settings, these systems complement dermatologists by providing second opinions or assisting in complex case diagnoses. A novel multi-task model showed improved diagnostic accuracy compared to traditional methods. However, the integration of AI in diagnosis raises ethical and practical concerns. Transparency and accountability are crucial, as AI models must be continuously monitored and validated to ensure their reliability in the dynamic field of skin diseases. These technologies can greatly enhance diagnostic accuracy, reduce physician burden, and improve patient outcomes, but successful implementation requires collaboration between technologists, clinicians, and policymakers to address ethical and practical challenges.

4.2.2. Potential Improvements and Extensions

The current research in automatic skin disease identification and auxiliary diagnosis has made notable advances, but optimization is still needed in several areas. For model improvement, optimizing both data quality and algorithm architecture is essential. Cross-center verification from the ISIC 2019 Challenge showed that training a model on high-definition dermatoscope images from a single device resulted in a 14.3% accuracy drop when tested with low-resolution images from smartphones. This emphasizes the need for multi-source datasets. The MDFNet team improved eczema and contact dermatitis differential diagnosis accuracy to 91.7% by integrating dermatoscope images with structured clinical data, offering new directions for clinical applicability. In feature extraction, a dual-attention network increased basal cell carcinoma recognition sensitivity to 89.5%, though false positives for pigmentation diseases remained at 16.2%, primarily due to skin tone differences. A study at Harvard Medical School found that when less than 15% of training data included dark-skinned samples, melanoma diagnostic specificity dropped by 23%. Hence, creating a demographically representative dataset is crucial. The FS3DCIoT framework's incremental learning strategy could dynamically optimize models across regions and races. To enhance diagnostic scalability, combining multi-dimensional information has proven effective. At Cleveland Medical Center, adding patient complaint data (e.g., itching degree, disease cycle) improved differential diagnosis of seborrheic dermatitis and psoriasis from 78.4% to 86.9%. However, synchronizing heterogeneous data remains challenging. Recent attempts with timing convolutional networks have performed well in chronic eczema prediction, offering a path for dynamic diagnostics. Transfer learning across diseases shows promise. Stanford's team applied a skin disease model to assist with COVID-19 CT imaging, achieving 93.2% accuracy. Transfer learning can also address rare skin diseases; fine-tuning a vitiligo model with 300 new samples increased recognition accuracy from 68% to 82%. However, care is needed regarding disease-specific characteristics to avoid spatial offsets during migration. Model compression and hardware adaptation are critical for clinical deployment. The enhanced MobileNet architecture reduced model size to 4.2MB while maintaining 91.3% accuracy, facilitating mobile deployment. Privacy protection remains an issue, though federated learning has proven effective, allowing multiple institutions to share knowledge without violating data privacy, as shown by a collaboration between 23 German clinics. Ethical considerations are also vital. In 2022, 63% of AI systems in skin disease diagnosis lacked sufficient decision-making transparency. Improvements using Class Activation Mapping (Grad-CAM) increased physician acceptance by 41%. A robust error traceability mechanism, such as the Mayo Clinic's diagnostic deviation analysis module, is necessary to ensure trust and facilitate system iterations. From an economic perspective, cost-effectiveness is key. A model from Cambridge University showed that AI systems with 85% diagnostic accuracy can reduce diagnosis costs by £34, provided daily treatments exceed 50 cases. This suggests tailored solutions for different healthcare levels: high-precision systems for Level 3 hospitals and lightweight tools for community clinics. Interdisciplinary collaboration is essential for overcoming technical barriers. The MHORUNet segmentation network, inspired by skin histopathology, achieved 82.1% in lesion boundary positioning. Expanding on bionic design ideas and leveraging immune cell recognition mechanisms can optimize model performance. Integration with Medical IoT technology, like wearable sensors for continuous psoriasis monitoring, offers new diagnostic possibilities.

A significant gap exists in standardization for technology transformation. The ISIC framework currently covers only 7 common skin tumors, leaving over 3,000 rare skin diseases underrepresented. Building a more fine-grained ontology library, as seen in Yonsei University's skin disease ontology mapping, could improve zero-sample recognition accuracy to 67.3%, expanding system capabilities. Future research should focus on creating a closed-loop intelligent system. A Mayo Clinic pilot project integrating AI diagnosis, treatment recommendations, and efficacy predictions shortened chronic urticaria diagnosis time from 6.2 weeks to 2.1 weeks, improving treatment efficiency by 28%. Multi-task learning frameworks that combine disease classification and severity assessment, like the model from Drexel University, offer a promising path forward, achieving an 87.9% weighted accuracy rate across 40 inflammatory skin diseases, a 11.6% improvement over single-task models.

4.2.3. Directions for Future Research

Future research in AI-driven skin disease diagnosis should focus on several key areas to enhance model capabilities. One promising direction is the exploration of advanced deep learning architectures, such as multimodal fusion methods like MDFNet, which combine skin images with clinical data to improve skin cancer classification. This approach integrates both visual and patient history data for more comprehensive diagnostics. Another critical area is leveraging big data and AI for personalized diagnosis. Large, diverse datasets combined with advanced algorithms can tailor models to individual patients, as demonstrated by ISIC 2019, where models trained on diverse datasets achieved over 90% accuracy in skin cancer diagnosis. However, challenges remain in ensuring these models are robust enough to account for variations in skin tones, ages, and ethnicities. Multi-center clinical research is essential to validate AI models across varied clinical scenarios. Studies across multiple dermatology clinics have shown that deep learning models can classify skin lesions accurately across different skin tones, highlighting their broader application potential. Explainable AI (XAI) is another crucial area. As AI models become more complex, transparency in decision-making is necessary. Techniques like dual attention networks can provide interpretable results, allowing dermatologists to better understand AI's recommendations and increase trust in the system. The integration of auxiliary tasks such as lesion segmentation, feature extraction, and patient history analysis can further enhance AI diagnostic support. A multi-task model developed for differentiating skin diseases not only classifies but also provides critical information, such as lesion boundaries and risk factors, aiding clinical decision-making. Personalized diagnosis, considering factors like age, gender, and genetics, is a promising area for AI. Models trained to incorporate these characteristics can improve diagnostic accuracy. For example, an enhanced MobileNet model improved actinic keratosis diagnosis by factoring in patient-specific features. Data augmentation and preprocessing are essential to overcome issues like poor resolution and variability in skin tones. Techniques such as rotation, scaling, and color normalization can improve model performance and generalization to new images. Collaborative research between AI developers and clinicians is vital. Close collaboration ensures that AI solutions address clinical challenges effectively. A study on contact dermatitis showed how integrating AI with clinical expertise can enhance diagnostic accuracy and support dermatologists in daily practice. Finally, the ethical implications of AI, such as data privacy, bias, and accountability, must be addressed. Bias in AI models, particularly when trained on non-representative datasets, is a significant concern. Ensuring fairness, transparency, and trust in AI systems is crucial for their successful integration into clinical practice.

In summary, future research should focus on advancing deep learning models, utilizing big data for personalized diagnosis, conducting multi-center studies, and integrating explainable AI. Developing multi-task models, improving data preprocessing, fostering collaboration, and addressing ethical concerns will be essential for the adoption and effectiveness of AI in clinical skin disease diagnosis.

Acknowledgments

This study was supported by the "Research on Automatic Recognition and Auxiliary Diagnosis of Artificial Intelligence in Skin Diseases" project of Guangxi College Students' Innovation and Entrepreneurship Training Program in 2024, project number: 202410599047.

References

[1] Li, Xuhui, et al. Image analysis and diagnosis of skin diseases-a review. Current Medical Imaging Reviews 19.3 (2023): 199-242.

[2] Gautam, Vinay, et al. Early Skin Disease Identification Using eep Neural Network. Comput. Syst. Sci. Eng. 44.3 (2023): 2259-2275.

[3] Anand, Vatsala, et al. An automated deep learning models for classification of skin disease using Dermoscopy images: A comprehensive study. Multimedia Tools and Applications 81.26 (2022): 37379-37401.

[4] Combalia, Marc, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. The Lancet Digital Health 4.5 (2022): e330-e339.

[5] Nigar, Natasha, et al. A deep learning approach based on explainable artificial intelligence for skin lesion classification. IEEE Access 10 (2022): 113715-113725.

[6] AlDera, Shaden Abdulaziz, and Mohamed Tahar Ben Othman. A model for classification and diagnosis of skin disease using machine learning and image processing techniques. International Journal of Advanced Computer Science and Applications 13.5 (2022).

[7] Wei, Zenghui, Qiang Li, and Hong Song. Dual attention-based network for skin lesion classification with auxiliary learning. Biomedical Signal Processing and Control 74 (2022): 103549.

[8] Chen, Qian, et al. MDFNet: application of multimodal fusion method based on skin image and clinical data to skin cancer classification. Journal of Cancer Research and Clinical Oncology 149.7 (2023): 3287-3299.

[9] Xiao, Junsheng, Jirui Li, and Honghao Gao. FS3DCIoT: A few-shot incremental learning network for skin disease differential diagnosis in the consumer IoT. IEEE Transactions on Consumer Electronics 69.4 (2023): 786-797.

[10] Groh, Matthew, et al. Deep learning-aided decision support for diagnosis of skin disease across skin tones. Nature Medicine 30.2 (2024): 573-583.

[11] Jones, O. T., et al. Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review. The Lancet Digital Health 4.6 (2022): e466-e476.

[12] Wu, Renkai, et al. Mhorunet: High-order spatial interaction unet for skin lesion segmentation. Biomedical Signal Processing and Control 88 (2024): 105517.

[13] Almuayqil, Saleh Naif, Sameh Abd El-Ghany, and Mohammed Elmogy. Computer-Aided diagnosis for early signs of skin diseases using multi types feature fusion based on a hybrid deep learning model. Electronics 11.23 (2022): 4009.

[14] Khurshid, Mahapara, Mayank Vatsa, and Richa Singh. Optimizing Skin Lesion Classification Via Multimodal Data and Auxiliary Task Integration. 2024 IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, 2024.

[15] Sun, Jie, et al. Machine learning methods in skin disease recognition: a systematic review. Processes 11.4 (2023): 1003.

[16] Agarwal, Raghav, and Deepthi Godavarthi. Skin disease classification using CNN algorithms. EAI Endorsed Transactions on Pervasive Health and Technology 9.1 (2023).

[17] Singh, Jagandeep, Jasminder Kaur Sandhu, and Yogesh Kumar. An analysis of detection and diagnosis of different classes of skin diseases using artificial intelligence-based learning approaches with hyper parameters. Archives of Computational Methods in Engineering 31.2 (2024): 1051-1078.

[18] Fraiwan, Mohammad, and Esraa Faouri. On the automatic detection and classification of skin cancer using deep transfer learning. Sensors 22.13 (2022): 4963.

[19] Escalé-Besa, Anna, et al. Exploring the potential of artificial intelligence in improving skin lesion diagnosis in primary care. Scientific Reports 13.1 (2023): 4293.

[20] Pesqué, David, et al. Differential diagnosis of contact dermatitis: a practical-approach review by the EADV Task Force on contact dermatitis. Journal of the European Academy of Dermatology and Venereology 38.9 (2024): 1704-1722.

[21] Kshirsagar, Pravin R., et al. Deep learning approaches for prognosis of automated skin disease. Life 12.3 (2022): 426.

[22] Anand, Vatsala, Sheifali Gupta, and Deepika Koundal. Skin disease diagnosis: challenges and opportunities. Proceedings of Second Doctoral Symposium on Computational Intelligence: DoSCI 2021. Springer Singapore, 2022.

[23] Ahammed, Mostafiz, Md Al Mamun, and Mohammad Shorif Uddin. A machine learning approach for skin disease detection and classification using image segmentation. Healthcare Analytics 2 (2022): 100122.

[24] Sangers, Tobias E., et al. Position statement of the EADV Artificial Intelligence (AI) Task Force on AI-assisted smartphone apps and web-based services for skin disease. Journal of the European Academy of Dermatology and Venereology 38.1 (2024): 22-30.

[25] Li, Zhouxiao, et al. Artificial intelligence in dermatology image analysis: current developments and future trends. Journal of clinical medicine 11.22 (2022): 6826.

[26] Inthiyaz, Syed, et al. Skin disease detection using deep learning. Advances in Engineering Software 175 (2023): 103361.

[27] Li, Shiyang, et al. An actinic keratosis auxiliary diagnosis method based on an enhanced MobileNet model. Bioengineering 10.6 (2023): 732.

[28] Lu, Jianyun, et al. Image classification and auxiliary diagnosis system for hyperpigmented skin diseases based on deep learning. Heliyon 9.9 (2023).

[29] Ye, Shengzhen, and Mingling Chen. The emerging role of artificial intelligence in diagnosis and clinical analysis of dermatology. Dermatologica Sinica 41.3 (2023): 145-152.

[30] Muhaba, Kedir Ali, et al. Automatic skin disease diagnosis using deep learning from clinical image and patient information. Skin health and disease 2.1 (2022): ski2-81.

[31] Parashar, Tarun, et al. Skin disease detection using deep learning. 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART). IEEE, 2022.

[32] Ghaffar Nia, Nafiseh, Erkan Kaplanoglu, and Ahad Nasab. Evaluation of artificial intelligence techniques in disease diagnosis and prediction. Discover Artificial Intelligence 3.1 (2023): 5.

[33] Li, Huanyu, et al. Deep skin diseases diagnostic system with Dual-channel Image and Extracted Text. Frontiers in Artificial Intelligence 6 (2023): 1213620.

[34] Li, Xiaoqiong, et al. Electret-based flexible pressure sensor for respiratory diseases auxiliary diagnosis system using machine learning technique. Nano Energy 114 (2023): 108652.

[35] Zhou, Yan-Jie, et al. A novel multi-task model imitating dermatologists for accurate differential diagnosis of skin diseases in clinical images. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2023.

[36] Vayadande, Kuldeep. Innovative approaches for skin disease identification in machine learning: A comprehensive study. Oral Oncology Reports (2024): 100365.

[37] Aruna R, Dr, et al. [Retracted] An Enhancement on Convolutional Artificial Intelligent Based Diagnosis for Skin Disease Using Nanotechnology Sensors. Computational Intelligence and Neuroscience 2022.1 (2022): 9539503.

[38] Zheng, Guiyue, et al. Research on skin disease health detection of college students based on deep learning. Journal of Physics: Conference Series. Vol. 2289. No. 1. IOP Publishing, 2022.

[39] Maduranga, M. W. P., and Dilshan Nandasena. Mobile-based skin disease diagnosis system using convolutional neural networks (CNN). International Journal of Image, Graphics and Signal Processing 12.3 (2022): 47.

[40] Ke, C. H. A. I., et al. Updated role of high-frequency ultrasound in assessing dermatological manifestations in autoimmune skin diseases. Acta Dermato-Venereologica 102 (2022): 1969.