Multi-Class Skin Disease Classification: A Study of Transfer Learning Strategies for Deep Learning Models

Paparao Mekala,
Department of CSE,
NIT Puducherry,
Karaikal, Puducherry, India.
paparao4u@gmail.com

B Surendiran,
Department of CSE,
NIT Puducherry,
*Karaikal,*Puducherry, India.
surendiran@gmail.com

P. Ramanathan,
Department of ECE,
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India,
pramanathan2509@gmail.com

K Ramanjaneyulu,
Department of CSE(DS),
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India,
20691A3233@mits.ac.in

S Sai Manikanta Prakash,
Department of CSE(DS),
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India,
20691A3238@mits.ac.in

T Reddy Venkata Chetan,
Department of CSE(DS),
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India,
20691A3236@mits.ac.in

N Abhiram varma
Department of CSE(DS),
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India,
20691A3201@mits.ac.in

***Abstract* - Skin diseases are a common health concern globally, requiring accurate diagnosis and treatment. Various techniques, including medication, topical treatments, phototherapy, and surgery, are employed for their cure. The detection of skin diseases has evolved significantly, with advancements in AI Technologies like image processing, machine learning, and deep learning. Image processing techniques extract meaningful features from images, while ML and DL algorithms classify diseases based on these features. In this work, we utilize a dataset of 23 classes, each containing 192 training images and 20 testing images. The dataset used here is the class-balanced dataset. We employ image processing techniques to enhance the dataset, followed by classification using DL models such as VGG19 and ResNet50. Our models achieve 83.69% and 82.82% accuracy, demonstrating their efficacy for classifying the skin disease images. This study contributes to the advancement of automation work of skin disease classification for diagnosis and treatment planning.**

Keywords - Skin diseases, dataset, image processing, CNN, ML, DL, classification, VGG19, ResNet50, DALYs, performance metrics.

I. INTRODUCTION

Skin disease is one of the common problems the majority of men face. A large portion of the global population is affected by skin diseases, impacting individuals of all ages and backgrounds. These conditions present a spectrum of severity, ranging from mild and temporary inconveniences to debilitating and life-threatening illnesses. This widespread prevalence underscores the critical need for advancements in diagnosis, treatment, and prevention [1]. An estimated 900 million people worldwide suffer from skin diseases at any given point, with five common conditions – fungal infections, pyoderma, scabies, acne, eczema, and warts – accounting for over 80% of all cases [2]. These conditions not only cause physical discomfort and social stigma but also significantly contribute to the global burden of disease, measured in Disability-Adjusted Life Years (DALYs) [1]. Research by Hay et al. (2014) [1] revealed that skin diseases contribute 1.79% to the global disease burden, highlighting their substantial impact on individual and public health.

However, the burden of skin diseases varies significantly across geographic regions and socioeconomic groups. We can see these issues where there is limited access and no exposure to cleanliness, healthcare, sanitation, and clean water. Especially in developing countries, it contributes to a higher prevalence of infectious skin diseases [3]. Other factors, such as age, lifestyle, etc., can influence a person's skin conditions.

Despite the widespread nature of skin diseases, advancements in medical research have led to the development of effective treatments for many prevalent conditions. For instance, topical and

systemic antibiotics have significantly improved the treatment of bacterial skin infections like pyoderma [4]. Similarly, topical corticosteroids and immunomodulators have proven effective in managing the symptoms of eczema and psoriasis [5]. For acne, a combination of topical medications, oral antibiotics, and hormonal therapies can offer significant relief [6]. Fungal infections, another common challenge, are often effectively addressed through the use of topical antifungals and oral medications [7]. While these advancements represent significant progress, the search for definitive cures and universally effective treatments continues. Research into the underlying causes of skin diseases, exploration of novel therapeutic options, and improved delivery systems remain crucial areas of focus.

Early and accurate diagnosis plays a critical role in effectively managing skin diseases. Traditionally, diagnosis has primarily relied on visual inspection by healthcare professionals. However, this approach can be subjective and prone to errors, particularly for complex or rare conditions. Also, the existing traditional methods are very expensive and complex in nature [8]. Technological advancements have introduced various innovative methods for detecting skin diseases. These techniques offer valuable tools for improved diagnostic accuracy and empower physicians to make informed treatment decisions.

The recent emergence of deep learning, a subfield of artificial intelligence, has sparked a revolution in the field of dermatology. Deep learning models are trained on various datasets of labeled images, enabling them to identify patterns and correlations that are even more difficult for the human eye to discern. These models shown promising results in the classification of various skin diseases, including melanoma, eczema, and psoriasis [9], [10], [11]. Studies by Nasr-Esfahani et al. (2018) [9] and Yu et al. (2017) [10] reported accuracies exceeding 90% in differentiating between malignant and benign lesions, demonstrating the immense potential of deep learning in this domain. Despite these impressive advances, several challenges remain. Ethical considerations regarding data bias, integration into existing healthcare workflows, and the need for continuous learning and adaptation require careful attention for the responsible and effective implementation of deep learning in dermatology. Alexnet, base model in DL Models [12].

Our study investigates the effectiveness of various deep-learning models for skin disease classification using the Skin Disease DermNet dataset. We employ two established architectures, VGG19 and ResNet50, and train them so that they can classify various skin diseases from the images. To assess the models' generalizability, we evaluate their performance on unseen data using accuracy, recall, precision, and F1-score metrics. This comparison will contribute to research advancements and potentially aid in developing practical tools for healthcare professionals.

II. LITERATURE SURVEY

Recent studies have delved into the efficacy of different deep-learning architectures for skin disease detection. Sapra et al. (2023) compared various CNN architectures on dermoscopic images, revealing that EfficientNet-B0 achieved an impressive accuracy of 92.3% while maintaining computational efficiency [13]. Additionally, Islam et al. (2023) explored the use of capsule networks, achieving an accuracy of 88.7% and demonstrating their potential for capturing spatial relationships within skin lesions [14]. These findings highlight the ongoing exploration of diverse architectures and their varying accuracies, showcasing the importance of selecting the most suitable architecture for the specific task.

While classification focuses on identifying the presence of a disease, early detection and accurate segmentation of lesions play crucial roles in improving patient outcomes. Studies like Yu et al. (2023) proposed a deep-learning model that classifies skin cancer more accurately and segments the cancerous region within the image, achieving impressive segmentation accuracy and providing valuable information for further diagnosis and treatment planning [15]. Similarly, Xie et al. (2023) investigated the use of deep learning for early detection of melanoma, achieving a promising accuracy of 87.5% in identifying suspicious lesions at an early stage [16]. These advancements show the potential of deep learning for early detection and segmentation, potentially leading to improved patient care and better treatment outcomes.

Data limitations are a critical challenge in utilizing deep learning for skin disease detection. The primary and foremost issue with the publicly accessible datasets is that skin lesions do not have a

enough data to train the models [17]. Datasets may be limited in size, imbalanced with overrepresentation of certain diseases, or lack diversity in terms of demographics or disease severity. These limitations can hinder model performance and lead to biased predictions. To address these challenges, studies like Wang et al. (2023) have proposed transfer learning techniques. By leveraging knowledge from pre-trained models on larger datasets, they achieved an accuracy of 89.2% on a smaller, specialized dataset, demonstrating the effectiveness of transfer learning in addressing data limitations [18] Additionally, Udupa et al. (2023) emphasize the importance of mitigating bias in healthcare AI by developing fairness-aware algorithms and ensuring diverse datasets that represent the broader population [19]. One of the common causes in medical image data analysis is insufficient data and imbalance [19].

While deep learning models demonstrate promising results in controlled settings, transitioning them into real-world clinical practice requires careful consideration. Studies like Tschandl et al. (2023) highlight the need for robust validation, regulatory approval, and clear guidelines for ethical use in clinical settings [20]. Effective integration necessitates collaboration between researchers, clinicians, and regulatory bodies to ensure responsible and trustworthy implementation.

The future of deep learning in skin disease detection and classification holds immense promise. Further research is needed to explore explainable AI techniques, ensuring transparency and building trust in model predictions. Additionally, continuous research into novel architectures, data augmentation strategies, and responsible AI practices is crucial to ensure ethical and equitable implementation of this technology across diverse healthcare settings.

III. METHODOLOGY

A. Architecture

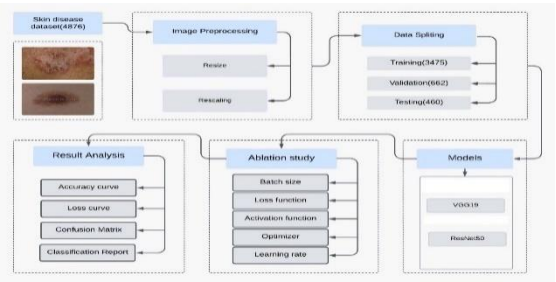

Fig. 1. Proposed Methodology Architecture

The model architecture begins with acquiring skin disease images, which are then preprocessed to enhance their quality and make them suitable for analysis. Next, a model, such as VGG19 or ResNet50, is constructed to classify skin diseases based on the preprocessed images. These models are well-trained using preprocessed images to learn the patterns and extract the features that distinguish different skin diseases. After training, the models are tested with a separate set of images called test datasets to determine their performance. In order to evaluate their performances, we used various evaluation metrics to assess how much the models can effectively classify skin diseases. This iterative process allows for the refinement of the model and the improvement of its accuracy in identifying various skin conditions.

B. Data Acquisition

This project utilizes the "Skin Disease DermNet" dataset available on Kaggle [12]. This Kaggle dataset provides valuable data for developers and researchers who are interested in and working in this field of skin disease classification. Comprising over 456MB of data, it encompasses a diverse range of 23 distinct skin disease classes.

The dataset offers a well-organized structure, divided into separate training and testing folders. Each class holds 192 images within the training set, allowing models to learn from substantial data. Additionally, 20 images per class are reserved in the testing set, enabling robust model performance evaluation on unseen data. Predominantly small, ranging from 50 to 150KB, the images are stored in the common RGB format, indicating three channels (Red, Green, and Blue) for capturing color information. This format is well-suited for deep learning models that rely on visual details for accurate classification.

The dataset delves into various skin conditions, encompassing common occurrences like "Acne and Rosacea" and more severe cases like " Basal Cell Carcinoma, Actinic Keratosis, and other Malignant Lesions." Additionally, it explores categories such as "Cellulitis Impetigo, "Bullous Disease," "Atopic Dermatitis," and other Bacterial Infections," showcasing the comprehensive coverage of diverse skin concerns. This level of diversity offers a significant advantage, as each class presents unique challenges for accurate identification. By training models on this dataset, researchers can ensure their

models are robust and capable of generalizing well to unseen data representing a wide spectrum of skin conditions.

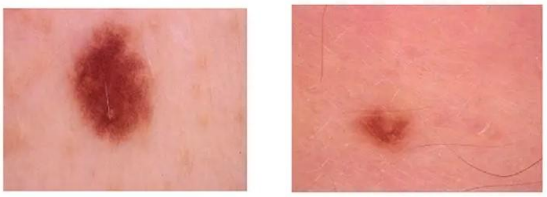

Fig. 2. Images after resizing

0 : Acne and Rosacea
1 : Actınıc Keratosıs Basal Cell Carcınoma and other Malıgnant Lesıons
2 : Atopıc Dermatıtıs
3 : Bullous Dısease
4 : Cellulıtıs Impetıgo and other Bacterıal Infectıons
5 : Eczema
6 : Exanthems and Drug Eruptıons
7 : Haır Loss Photos Alopecıa and other Haır Dıseases
8 : Herpes HPV and other STDs
9 : Lıght Dıseases and Dısorders of Pıgmentatıon
10 : Lupus and other Connectıve Tıssue Dıseases
11 : Melanoma Skın Cancer Nevı and Moles
12 : Naıl Fungus and other Naıl Dısease
13 : Poıson Ivy Photos and other Contact Dermatıtıs
14 : Psorıasıs pıctures Lıchen Planus and Related Dıseases
15 : Scabıes Lyme Dısease and other Infestatıons and Bıtes
16 : Seborrheıc Keratoses and other Benıgn Tumors|
17 : Systemıc Dısease
18 : Tınea Rıngworm Candıdıasıs and other Fungal Infectıons
19 : Urtıcarıa Hıves
20 : Vascular Tumors
21 : Vasculıtıs Photos
22 : Warts Molluscum and other Vıral Infectıons

Fig. 3. Class names and its labels

C. Data Preprocessing

The imported dataset contains two directories i.e., train and test datasets. The training dataset is used to train the model. It contains images from each class of skin disease, with 192 images per class. The validation dataset is used to fine-tune the model's hyperparameters so that the model can prevent overfitting. It is created by randomly selecting 15% of the training dataset for each class. The test dataset evaluates and estimates the model's performance on unseen data. It contains 20 images per class, separate from the training and validation dataset images.

The first crucial step in training a deep learning model is ensuring all images adhere to a consistent size. This uniformity allows the model to process information efficiently and avoid biases resulting from varying pixel dimensions. A process was described to transform the dataset for the skin disease classification model. Each folder in every directory represented a distinct skin disease class. The process involved randomly sampling a subset of images from each class to ensure a balanced representation.

These images are resized to a fixed size of 224x224 pixels and added to each individual defined array along with their corresponding class labels in the form of tuples. This methodology aimed to create a comprehensive dataset with an even distribution of images across all classes, enhancing the model's ability to classify various skin diseases accurately.

After resizing, the data is prepared to train a classification model on image data. Class labels are converted into numerical form and then further transformed into a format suitable for the model. The input images are also preprocessed to ensure they are in a standardized format for training. These steps collectively enable the data to be used effectively for training the classification model.

D. Model Building

After Data preprocessing, transfer learning is applied to two widely used Convolutional Neural Network (CNN) Architectures - VGG19 and ResNet50 to develop a skin disease classification model. Using transfer learning techniques, we will fine-tune the models with the pre-trained weights that are already trained for specific tasks on standard datasets. This method is advantageous as it allows us to leverage the learned features from the pre-trained models, which excel at capturing generic image features like edges, shapes, and textures.

The VGG19 model, created by the Visual Geometry Group at the University of Oxford. Due to its nature of simplicity and effectiveness, we have chosen this model. The model is comprised of 19 layers of different convolutional layers, 3x3 filters, and max-pooling layers. By fine-tuning the VGG19 model on our skin disease dataset, we aim to utilize its ability to extract detailed features from images. The ResNet50 model, developed by Microsoft Research, was selected for its deep architecture with residual connections. These connections enable training very deep networks (50 layers) without issues like vanishing gradients. Fine-tuning ResNet50 on our dataset aims to benefit from its depth and capability to capture intricate image patterns.

This transfer-learning approach for skin disease classification uses models like VGG19 and

ResNet50. The process involves initializing the base model with weights pre-trained on ImageNet and excluding the top layers to retain the pre-trained features. These base models are then frozen to prevent weight updates during training. Additional layers are added to the base models, including global average pooling and dense layers for feature extraction and classification. The output layer consists of units equal to the number of classes (23 in our case), and the activation function we used here for multi-class classification is SoftMax. Compiling the model involves selecting an optimizer like Adam and setting a learning rate (e.g., lr=0.001), and we used the loss function here is categorical cross-entropy. The learning rate is very important and determines the step size selected during optimization, influencing the speed and quality of model convergence. The categorical cross-entropy measures the difference between predicted and actual class distributions, guiding the model to minimize this difference during training. These components are essential for training effective skin disease classification models, ensuring the models learn meaningful features and make accurate predictions.

E. Model Training and Testing

The compiled models are well-trained using the preprocessed training dataset during the training phase. This process involves optimizing the model's weights to minimize the loss, which in this case is categorical cross-entropy. Each model undergoes training over a specific number of epochs and with a predefined batch size.

As the training progresses, the model learns to extract the most relevant features from input images and associate them with the correct skin disease classes. The validation dataset monitors the training progress and evaluates the model's performance on unseen data, which helps prevent overfitting. Overfitting occurs when the model performs well on the training data but fails to generalize to new data. Hyperparameters such as the learning rate, batch size, and number of epochs are adjusted during training to find the optimal set of weights for each model. Regularization techniques like dropout or weight decay are also applied to prevent overfitting and enhance the model's ability to generalize.

Each trained skin disease classification model is assessed using a separate unseen test dataset in the testing phase. The model predicts the classes of the test images, and these predictions are compared with the actual labels to determine the model's accuracy. Testing is crucial for assessing the trained model and how well the model generalizes to the unseen data. It also helps identify issues like overfitting or underfitting, where the model performs well on the training data but poorly on unseen data.

F. Evaluation Metrics

In evaluating the effectiveness of skin disease classification models, performance metrics used in our work are accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the models' performance, considering both the correctness and the ability to detect relevant instances.

Accuracy: It provides a general overview of the model's correctness and is widely used in classification tasks.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Prediction}} \quad (1)$$

Precision: Precision is particularly important in medical applications, including skin disease classification, where correctly identifying positive cases is crucial to avoid misdiagnosis.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}} \quad (2)$$

Recall (Sensitivity): It indicates the proportion of correctly predicted positive instances out of all actual positive instances. Recall is crucial in medical diagnostics as it ensures that all instances of the disease are correctly identified, minimizing false negatives.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}} \quad (3)$$

F1-score: The F1-score is the harmonic mean of precision and recall. It is particularly useful when there is an uneven class distribution.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}} \quad (4)$$

These metrics will play an important role in evaluating the model's performance, aiding in refining the better model. Different models of state-of-the-art models can be compared with our models for the same work of image classification[21]. Some automated models have also been implemented for current research work [22]

IV. RESULTS AND DISCUSSION

In the result analysis, the evaluation process involves assessing the skin disease classification models using metrics like accuracy, precision, recall, and F1-score. Confusion matrices are also used to understand the models' performance for each disease class.

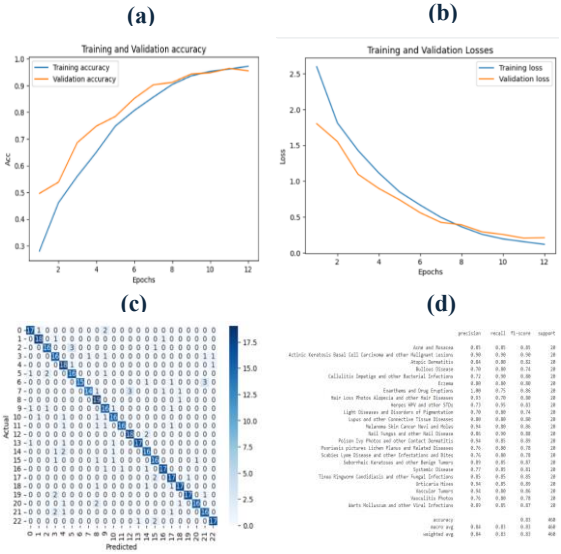

Fig. 4. Summary of ResNet50 model (a) Training and Validation Accuracy (b) Training and Validation Loss (c) Confusion matrix (d) Classification report

TABLE 1. PERFORMANCE OF RESNET50 MODEL

Epochs	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
3/12	0.5863	0.6651	1.4100	1.0730
6/12	0.8238	0.8816	0.6164	0.5234
9/12	0.9452	0.9415	0.2526	0.2998
12/12	0.9701	0.9610	0.1309	0.2071

In the examination of skin disease classification employing the ResNet50 model, a meticulous scrutiny unfolded. The training loss exhibited a discernible descent from 2.59 to 0.119, and the validation loss mirrored this trend, declining from 1.802 to 0.211 over 12 epochs and batch size 32. This nuanced loss dynamics underscore the model's adeptness in iterative training, progressively refining its predictive capacity. Concurrently, accuracy metrics unveiled noteworthy improvements, with training accuracy ascending from 0.28 to an impressive 0.9719, and validation accuracy demonstrating a robust rise from 0.496 to 0.955 across the 12 epochs, affirming the model's capacity for effective generalization. An examination of the confusion matrix revealed precise class predictions, with isolated instances of misclassification involving 4 or 5 images in specific classes. This nuanced misclassification indicates the model's resilience to outliers and its ability to adapt to diverse instances within the dataset. Precision values spanning from 70% to 94% underscore the model's refined discriminative capabilities. Concurrently, recall values, predominantly oscillating between 85%, 80%, or 95%, highlight the model's sensitivity to effectively capturing instances of each class. The weighted average accuracy of 84%, weighted recall accuracy of 83%, and weighted average F1 score of 83% collectively emphasize the model's robust and consistent performance in categorizing diverse skin diseases across multiple classes.

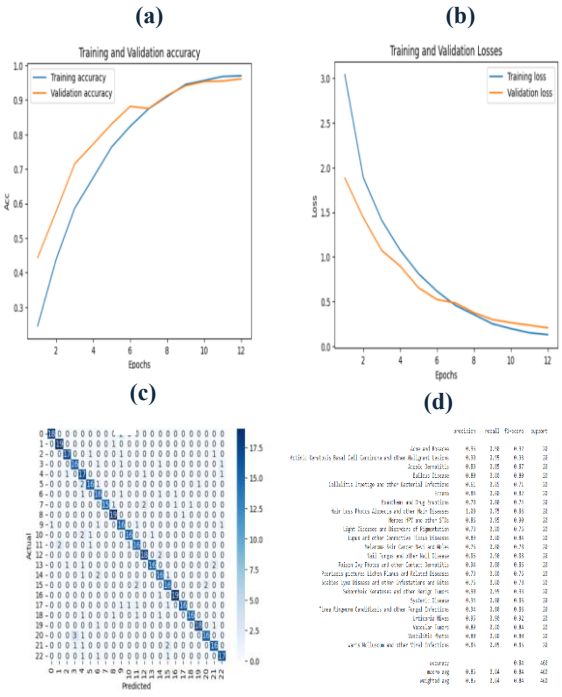

Fig. 5. Summary of VGG19 model (a) Training and Validation Accuracy (b) Training and Validation Loss (c) Confusion matrix (d) Classification report

TABLE 2. PERFORMANCE OF VGG19 MODEL

Epochs	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
3/12	0.5607	0.6167	1.4278	1.0956
6/12	0.8066	0.8516	0.6669	0.5602
9/12	0.9361	0.9330	0.2579	0.2926
12/12	0.9719	0.9550	0.1190	0.2112

In scrutinizing skin disease classification with the VGG19 model, a 12-epoch and batch size 32 training regimen revealed its adept learning

capabilities. The training loss descended from 3.03 to 0.13, mirroring a similar trend in validation loss from 1.88 to 0.207. Accompanying this, training accuracy climbed commendably from 0.245 to an outstanding 0.9701, and validation accuracy improved robustly from 0.443 to 0.961. The confusion matrix highlighted precise class predictions with minimal misclassifications, typically involving 2 to 4 images within specific classes. Precision values ranged from 70% to 95%, while recall values clustered around 75% to 95%, and F1 scores maintained a balance between 74% and 93%. The model exhibited a weighted average accuracy of 85%, emphasizing its proficiency across all classes. The weighted recall accuracy stood at 84%, and the weighted average F1 score registered at 84%, attesting to the VGG19 model's consistent and balanced performance in multi-class skin disease classification.

TABLE 3. PERFORMANCE OF CNN MODELS

Model Name	Training Accuracy	Validation Accuracy	Test Accuracy
Resnet50	97.19	95.50	82.82
VGG19	97.01	96.10	83.69

The table summarizes the performance of ResNet50, VGG19 models in classifying skin diseases. VGG19 achieved the highest test accuracy of 83.60%, outperforming the other models. These results suggest that VGG19's efficient architecture allows it to extract relevant features effectively from skin disease images, making it a suitable choice for resource-constrained environments.

V. CONCLUSION

In this study, we investigated the performance of two deep learning models, namely VGG19, and ResNet50, for the classification of skin diseases. Our findings reveal that VGG19 exhibited the highest accuracy than the ResNet50 model, achieving a test accuracy of 83.69%. This indicates the potential of VGG19 as an effective tool for the classification of skin diseases, which is crucial for accurate diagnosis and timely treatment. Comparatively, ResNet50 achieved a test accuracy of 82.82%, also showed respectable performance, superior accuracy highlights its suitability for practical applications in dermatology. The higher accuracy of can be attributed to its architecture, which is optimized for efficiency and performance. Overall, our study demonstrates the effectiveness of deep learning models in the classification of skin diseases. These models hold great promise for improving the accuracy and efficiency of skin disease diagnosis, ultimately benefiting patients and healthcare providers alike. Further research and development in this area could lead to the development of more advanced and accurate diagnostic tools for dermatological conditions.

REFERENCES

[1] R. J. Hay *et al.*, “The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions,” *J Invest Dermatol*, vol. 134, no. 6, pp. 1527–1534, 2014, doi: 10.1038/JID.2013.446.

[2] “Skin diseases.” Accessed: Mar. 03, 2024. [Online]. Available: https://platform.who.int/mortality/themes/theme-details/topics/topic-details/MDB/skin-diseases

[3] R. J. Hay *et al.*, “The Global Burden of Skin Disease in 2010: An Analysis of the Prevalence and Impact of Skin Conditions,” *Journal of Investigative Dermatology*, vol. 134, pp. 1527–1534, 2014, doi: 10.1038/jid.2013.446.

[4] “Eczema (Atopic Dermatitis) | NIAID: National Institute of Allergy and Infectious Diseases.” Accessed: Mar. 03, 2024. [Online]. Available: https://www.niaid.nih.gov/diseases-conditions/eczema-atopic-dermatitis

[5] “Psoriasis | NIAMS Catalog.” Accessed: Mar. 03, 2024. [Online]. Available: https://catalog.niams.nih.gov/health-topics/psoriasis

[6] “Fungal Diseases | NIAID: National Institute of Allergy and Infectious Diseases.” Accessed: Mar. 03, 2024. [Online]. Available: https://www.niaid.nih.gov/diseases-conditions/fungal-diseases

[7] D. Wang *et al.*, “Modeling early treatment response in AML from cell-free tumor DNA,” *iScience*, vol. 26, no. 12, p. 108271, Dec. 2023, doi: 10.1016/J.ISCI.2023.108271.

[8] P. Mekala, S. B, G. Vishnu, G. J. Sai, and P. Subash, “Classification of Skin Lesions using Deep Learning Models with Transfer Learning Techniques,” in *2023 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE)*, 2023, pp. 1–10. doi: 10.1109/RASSE60029.2023.10363548.

[9] J. Ker, L. Wang, J. Rao, and T. Lim, “Deep Learning Applications in Medical Image Analysis,” *IEEE Access*, vol. 6, pp. 9375–9379, Dec. 2017, doi: 10.1109/ACCESS.2017.2788044.

[10] L. Yu, H. Chen, Q. Dou, J. Qin, and P. A. Heng, “Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks,” *IEEE Trans*

Med Imaging, vol. 36, no. 4, pp. 994–1004, Apr. 2017, doi: 10.1109/TMI.2016.2642839.

[11] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature 2017 542:7639*, vol. 542, no. 7639, pp. 115–118, Jan. 2017, doi: 10.1038/nature21056.

[12] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," *PLoS One*, vol. 14, no. 5, May 2019, doi: 10.1371/journal.pone.0217293.

[13] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature 2017 542:7639*, vol. 542, no. 7639, pp. 115–118, Jan. 2017, doi: 10.1038/nature21056.

[14] M. Dash, N. D. Londhe, S. Ghosh, A. Semwal, and R. S. Sonawane, "PsLSNet: Automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network," *Biomed Signal Process Control*, vol. 52, pp. 226–237, Jul. 2019, doi: 10.1016/J.BSPC.2019.04.002.

[15] H. Nahata and S. P. Singh, "Deep Learning Solutions for Skin Cancer Detection and Diagnosis," *Learning and Analytics in Intelligent Systems*, vol. 13, pp. 159–182, 2020, doi: 10.1007/978-3-030-40850-3_8/COVER.

[16] P. Lambin *et al.*, "Radiomics: Extracting more information from medical images using advanced feature analysis," *Eur J Cancer*, vol. 48, no. 4, pp. 441–446, Mar. 2012, doi: 10.1016/J.EJCA.2011.11.036.

[17] P. Yao *et al.*, "Single Model Deep Learning on Imbalanced Small Datasets for Skin Lesion Classification." [Online]. Available: https://github.com/yaopengUSTC/mbit-skin-cancer.git.

[18] D. A. Okuboyejo and O. O. Olugbara, "A Review of Prevalent Methods for Automatic Skin Lesion Diagnosis," *Open Dermatol J*, vol. 12, no. 1, pp. 14–53, May 2018, doi: 10.2174/187437220181201014.

[19] L. Wang, L. Zhang, X. Shu, and Z. Yi, "Intra-class consistency and inter-class discrimination feature learning for automatic skin lesion classification," *Med Image Anal*, vol. 85, p. 102746, Apr. 2023, doi: 10.1016/J.MEDIA.2023.102746.

[20] B. Wan *et al.*, "Applications and future directions for optical coherence tomography in dermatology," *British Journal of Dermatology*, vol. 184, no. 6, pp. 1014–1022, Jun. 2021, doi: 10.1111/BJD.19553.

[21] M. Rahaman, M. Grzegorzek, P. Zhao, C. Li, H. Yang, and T. Jiang, "A Comparison of Deep Learning Classification Methods on Small-scale Image Data set: from Converlutional Neural Networks to Visual Transformers A Comparison of Deep Learning Classification Methods on Small-scale Image Data set: from Convolutional Neural Networks to Visual Transformers," 2021. [Online]. Available: https://www.researchgate.net/publication/353330756

[22] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," *IEEE Trans Med Imaging*, vol. 36, no. 4, pp. 994–1004, 2017, doi: 10.1109/TMI.2016.2642839.