Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

RESEARCH ARTICLE OPEN ACCESS

# Deep Learning-Based Detection of Hair and Scalp Diseases Using CNN and Image Processing

**Busireddy Seshakagari Haranadha Reddy**

Manager – Architecture, Valuemomentum, Erie, PA, USA 16506.

DOI: **10.5281/zenodo.14965660**

Received: 13 February 2025 / Revised: 28 February 2025 / Accepted: 04 March 2025


**Abstract –** Hair and scalp disorders affect millions worldwide, often undiagnosed in their early stages due to similarities with normal hair shedding. Conditions such as alopecia, psoriasis, and folliculitis require professional dermatological evaluation, which can be time-consuming and lead to delayed treatment. Early diagnosis and clinical decision-making have been strengthened by automating disease detection in healthcare through the development of deep learning and improved image-processing techniques. This study evaluated the three primary conditions associated with the scalp using a convolutional neural network (CNN) model. The research faced challenges due to the restricted availability of structured datasets and variations in image quality from several sources. To address this, 150 images were gathered from various repositories and subjected to pre-processing techniques, including denoising, contrast enhancement, image equalization, and data balancing. The processed dataset was employed to train a 2D CNN model with a training and validation accuracy of 97% and 92%. Additionally, we curated a dataset of scalp images to support further research in this domain, promoting the development of more robust automated diagnostic systems.

**Index Terms** – Hair and Scalp Disorders, Deep Learning, Medical Image Processing, Convolutional Neural Networks (CNNs), Image Pre-processing Techniques, Skin and Scalp Analysis.

## I. INTRODUCTION

Millions of people throughout the world suffer from hair and scalp diseases, which can cause profound emotional and physical suffering [1]. Psoriasis, folliculitis, and alopecia are some of the most prevalent

MILESTONE RESEARCH.IN OPEN ACCESS
Electronic ISSN: **2584-072X**

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

conditions affecting the scalp [2]. However, the symptoms' mild beginning and resemblance to regular hair shedding make early detection difficult. Many patients struggle to differentiate between temporary hair loss caused by stress or seasonal changes and pathological conditions requiring medical intervention. Traditional diagnosis methods rely heavily on dermatological expertise, where specialists perform visual examinations and, in some cases, medical tests to confirm the presence of a disease [3]. However, people in distant locations may not always have access to such diagnostic methods, which are costly and time-consuming [4]. Therefore, if discovery and treatment are delayed, the condition may worsen and cause permanent hair loss and serious scalp injury. Scholars are exploring the possibilities of deep learning (DL) and artificial intelligence (AI) in healthcare applications due to the growing need for automated and effective diagnostic tools [5].

In recent years, DL, particularly convolutional neural networks (CNNs), has revolutionized medical image analysis by providing automated solutions for detecting and classifying diseases with high precision [6]. By extracting complex characteristics from medical pictures, CNNs are widely used to diagnose dermatological disorders, including vitiligo, eczema, and melanoma [7]. Applications in healthcare and diagnostic medicine are well suited for DL models because of their capacity to learn from massive datasets and spot patterns that are invisible to the human eye. Although AI-driven dermatology has made great strides, nothing is known about using DL to identify scalp diseases. Scalp disorders differ from general skin problems, frequently appearing as separate lesions or aberrant pigmentation. This is because hair obscures the afflicted areas and complicates image processing. The absence of organized and openly accessible datasets is one of the main obstacles impeding the creation of DL models for scalp disease identification. Although there are extensive databases for diseases like skin cancer, no specific dataset has been assembled for problems with the scalp. Because of this, researchers are forced to rely on photographs gathered from various sources, such as internet repositories and medical literature, which differ in image quality, illumination, and resolution. These discrepancies hamper practical DL model training because misclassification and significant false-positive or false-negative rates might result from differences in picture characteristics. Moreover, scalp disorders can spread over the scalp, eyebrows, eyelashes, beard, and other hair-bearing regions, making segmentation and localization even more difficult. The unequal distribution of hair among persons presents another challenge since illnesses may be more difficult to detect in those with thick or curly hair. On the other hand, lighter or sparser hair may expose more scalp surfaces, leading to biased detection outcomes.

To address these challenges, this study proposes a CNN-based DL model capable of detecting and classifying scalp diseases, including alopecia, folliculitis, and psoriasis. Owing to the lack of a consistent dataset, we compiled an image dataset by sourcing 150 scalp images from various online medical platforms and dermatological sources. Large-scale preprocessing was performed on these photos, including image equalization, scaling, contrast enhancement, and denoising, to increase quality and ensure consistency throughout model training. After the images were processed, they were put into a 2D CNN architecture, which effectively and accurately categorized the three illnesses. The motivation behind this research derives from the critical need for accessible, automated, and reliable disease detection systems. An essential part of avoiding this is early progression of the disease and improving treatment results, but dependence on traditional methods usually results in incorrect delays and diagnosis. This research aims to close the gap between AI-driven dermatology and scalp disease diagnosis by utilizing the potential of DL

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

by giving patients a preliminary evaluation of their condition and developing a strong, image-based categorization system that improves clinical decision-making and offers patients more control. These tools have the potential to significantly lessen the workload for dermatologists and increase access to treatment globally, provided AI continues to progress and databases on scalp diseases continue to grow.

## II. LITERATURE SURVEY

This section will discuss several existing works that use machine learning (ML) and DL to revolutionize disease detection in health informatics, particularly dermatological applications, where image-based analysis enables rapid and precise diagnosis. Recent improvements have enabled the automation of disease identification, minimized human errors, and expedited clinical decision-making. Chowdhury et al. [8] leveraged a modified Xception model, integrating ReLU activation, dense layers, global average pooling, regularization, and dropout layers to enhance its performance in diagnosing hair and scalp disorders. This DL approach is systematically evaluated against well-established models such as VGG19, Inception, ResNet, and DenseNet to assess its effectiveness in accurate classification. Outperforming the comparison models, which had accuracy levels between 50% and 80%, the updated Xception model performed exceptionally well, with an astounding 92% accuracy rate. To ensure transparency and interpretability in the model's predictions, Explainable AI (XAI) techniques were employed, including Gradient-weighted Class Activation Mapping (Grad-CAM) and Saliency Map.

Shakeel et al. [9] proposed a framework to differentiate alopecia areata from healthy hair using image-based analysis by collecting 200 healthy hair images from the Figaro1K dataset and 68 alopecia areata images from DermNet. With a 10-fold cross-validation approach, Support Vector Machines (SVM) and K-nearest neighbors (KNN) achieved accuracy rates of 91.4% and 88.9%, respectively, in the classification challenge. Notwithstanding these encouraging outcomes, the study had several drawbacks. First, other ML methods that would have increased categorization accuracy were not investigated. Kapoor et al. [10] developed an ML-based model for early alopecia detection using a neural network approach utilizing 100 samples, with 80% allocated for training and 20% for testing and focused on four attributes: hair length, follicle health, nail brittleness, and hair damage. A two-layer feed-forward neural network with backpropagation was employed for classification, consisting of four input neurons, ten hidden neurons, and a linear output neuron. The model achieved a training accuracy of 91% and a validation accuracy of 86.7%, demonstrating its effectiveness. The study reported optimal performance at epoch four with a gradient of 0.059687.

In [11], an ML model was developed to classify eczema, melanoma, and psoriasis. The dataset comprised 80 images collected from various dermatological sources. Feature extraction employed a convolutional neural network (CNN) and classification using a multiclass SVM. The model achieved an impressive accuracy of 100%. Another study [12] explored the classification of skin lesions into five categories, healthy, acne, eczema, benign, and malignant melanoma, by employing a pre-trained CNN model, AlexNet, for feature extraction and using an error-correcting output code for classification. The dataset included 9,144 images sourced from multiple dermatological repositories. The model achieved an accuracy of 84.21% using a 10-fold cross-validation technique. Although ML has been extensively studied for skin disease classification, few studies have explicitly focused on scalp diseases. Hair-covered scalp

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

regions pose additional challenges for image-based disease classification due to occlusion, making segmentation and feature extraction more complex. Many current research either fail to address model dependability or don't consider inter-class differences, which might result in misclassification mistakes. In addition, overfitting and failing to assess essential performance indicators are still common in the literature today [16] ,[17]. By putting into practice, a CNN-based architecture for scalp illness identification, our study fills these gaps while guaranteeing excellent precision, recall, and overall classification accuracy. Our strategy focuses on increasing dataset variety, refining preprocessing methods, and adding strong validation measures to improve model performance and dependability [18]

## III. METHODS & MATERIALS

This section outlines the workflow of our proposed model and details each module's role. As depicted in Fig. 1, the procedure starts with image preprocessing, and the refined images are provided into a convolutional neural network (CNN) for classification. The model is designed to accurately categorize images into three distinct conditions: alopecia, folliculitis, and psoriasis.

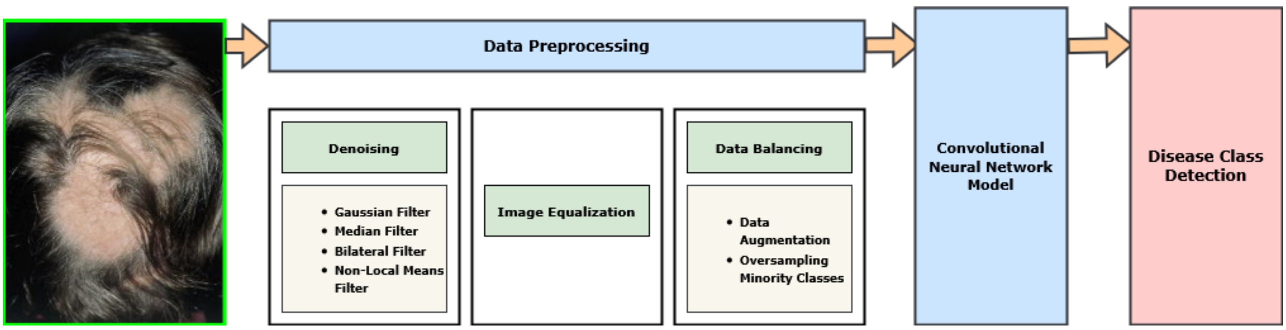


**Fig. 1:** Graphical Representation of the Overall Methodology

### *A. Dataset Description*

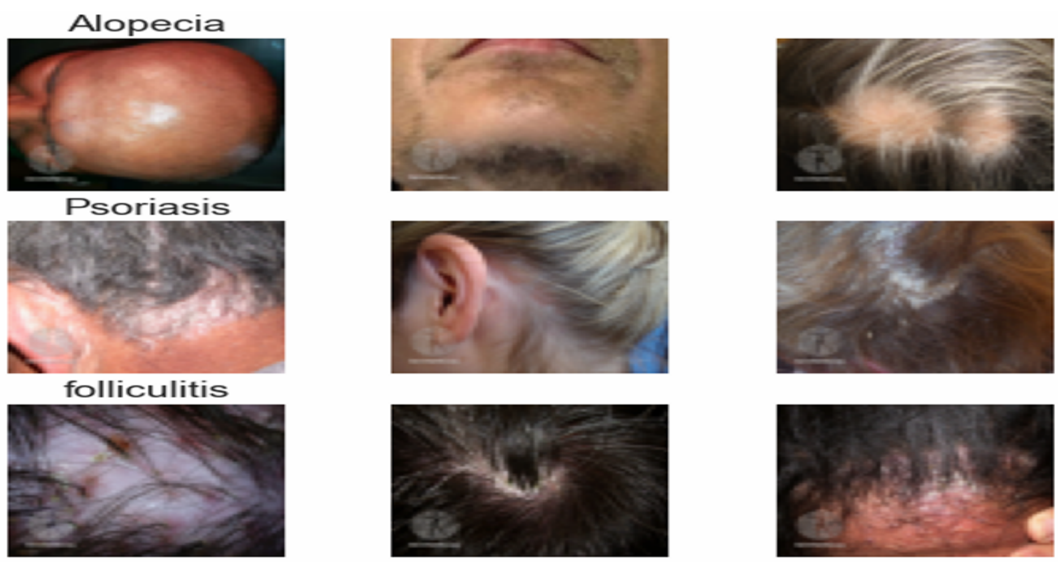


**Fig. 2:** Sample of the Dataset

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

Relevant and high-quality images for specific hair and scalp conditions are often limited and scattered across online sources. To address this, we employed an exhaustive dataset by sourcing images from multiple reputable dermatology platforms, including DermQuest, DermNet, MedicineNet, and DermnetNZ, as well as contributions from medical professionals [13]. The distribution of images in various disorders varies greatly, with alopecia cases being more common because of the nature and general severity of the condition. Consequently, alopecia photographs are included in the database in more significant quantities than images of other scalp conditions. This dataset contains a total of 150 images of Alopecia at 65, Psoriasis at 45, and Folliculitis at 40. The entire number of samples for each category is detailed in Table I, while a visual representation of randomly selected photos from each class is provided in Fig. 2.

**TABLE I:** Number of Images with Diseases

| Disease | Number of images |
|---|---|
| Alopecia | 65 |
| Psoriasis | 45 |
| Folliculitis | 40 |
| Total | 150 |

*B. Experimental Setup and Hardware Configuration*

For this purpose, we utilized an HP EliteBook 840 13th-generation Intel Core i5 laptop running Windows 11. The device offered effective processing and storage capabilities with a 256 GB M.2 PCIe NVMe SSD and 16 GB DDR4 RAM (3200 MHz). Image classification tasks were performed using the integrated Intel Iris Xe graphics, which supports Thunderbolt technology and enhances computational efficiency for deep learning applications. iPhone 13 Pro Max was used for data collection which features a Hexa-core processor (2x3.23 GHz Avalanche + 4x1.82 GHz Blizzard) and an Apple GPU with 5-core graphics. The mobile device's 128 GB storage and 6 GB RAM gave it enough space to handle images with a high resolution. A 12-megapixel primary camera with three lenses was used to take detailed pictures to ensure high-quality input data for the classification process.

*C. Data Pre-processing*

Data preparation is necessary to offer a balanced representation in different classes, improving model performance and picture quality. Our preprocessing procedure has three main parts: denial, image equalization, and data balance.

*i. Denoising*

Images captured frequently have noise from outside sources that causes erratic changes in color and brightness. Our dataset's non-uniform noise distribution, including photos gathered from several dermatological sources, presents further image processing difficulties. To solve this, a variety of filtering strategies were examined.

Milestone Transactions on
**MEDICAL TECHNOMETRICS**
Vol. 03, Issue. 01, 2025

- *Gaussian Filter:* Initially applied for noise reduction but resulted in excessive blurring, causing loss of critical image details and edge degradation.
- *Median Filter:* Performed better than the Gaussian filter, particularly in reducing impulse noise, but lacked optimal edge preservation.
- *Bilateral Filter:* Showed improved noise reduction while maintaining edge sharpness but introduced minor artifacts in some cases.
- *Non-Local Means Filter***:** Provided the best results by effectively reducing noise while preserving fine details and edges, making it the optimal choice for our application. The effectiveness of this filter is illustrated in Fig. 3.

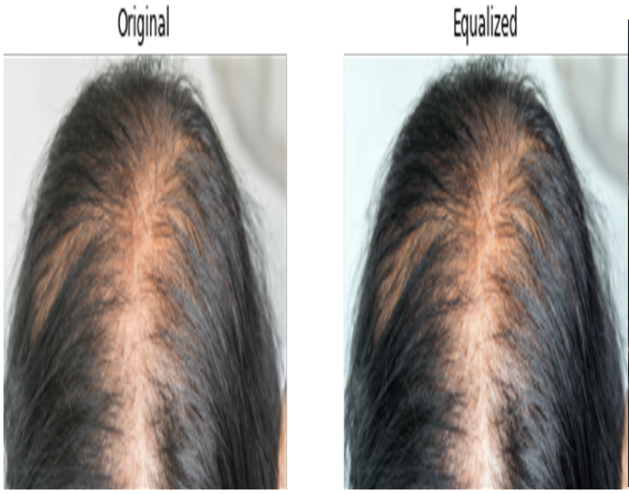


**Fig. 3:** Original vs. Non-Local Means Denoised

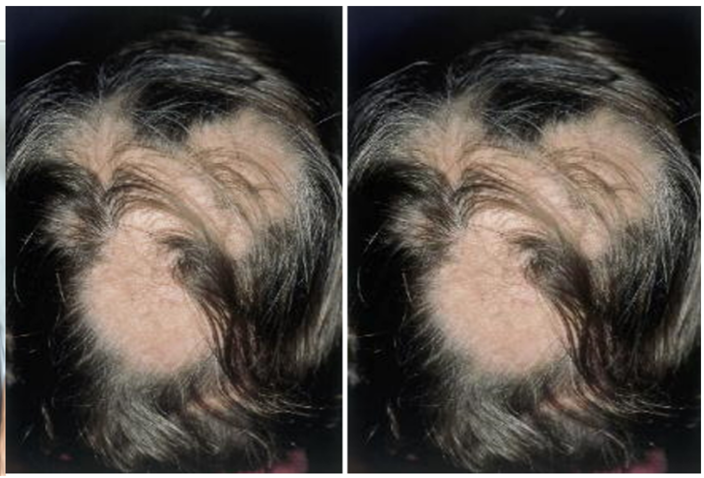

**Fig. 4:** Image equalization utilizing CLAHE

ii. *Image Equalization*

Raw photos must be enhanced to ensure a more accurate and consistent depiction because they frequently suffer from irregular contrast and brightness changes. We first used Histogram Equalization (HE) to improve contrast. However, this method arbitrarily changed global contrast, which occasionally resulted in excessive background amplification and the loss of localized features. To circumvent these restrictions, contrast-limited adaptive histogram equalization (CLAHE) was employed. This technique ensures that contrast modifications do not cause over-amplification by performing localized contrast enhancement and dividing the picture into smaller, non-overlapping sections. By efficiently normalizing photos, CLAHE enhances visibility without sacrificing fine details. Figure 4 presents the impact of CLAHE on picture equalization.

iii. *Data Balancing*

A well-balanced dataset is essential to avoid model bias toward the majority class. The categorization findings may be biased because our dataset had many more pictures of alopecia than other scalp disorders. To address this imbalance, we combined oversampling and data augmentation methods:

- *Data Augmentation:* Introduced manipulated images using cropping, random rotation, rescaling, and flipping them vertically and horizontally, which directed to better feature representation and increased model generalization.

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

- *Oversampling Minority Classes:* Synthetic samples were developed to expand the representation of underrepresented conditions, providing the model learns evenly from all categories.

*C. Neural Network Model*

Neural networks are essential in visual data analysis because they allow for automatic feature extraction and the identification of intricate patterns [14]. Among many designs, CNNs do very well in image-based classification applications [15]. In this investigation, we used an AutoKeras-optimized CNN model, which investigated 25 possible configurations before deciding on the most effective design. Fig. 5 illustrates the architecture of the neural network model.

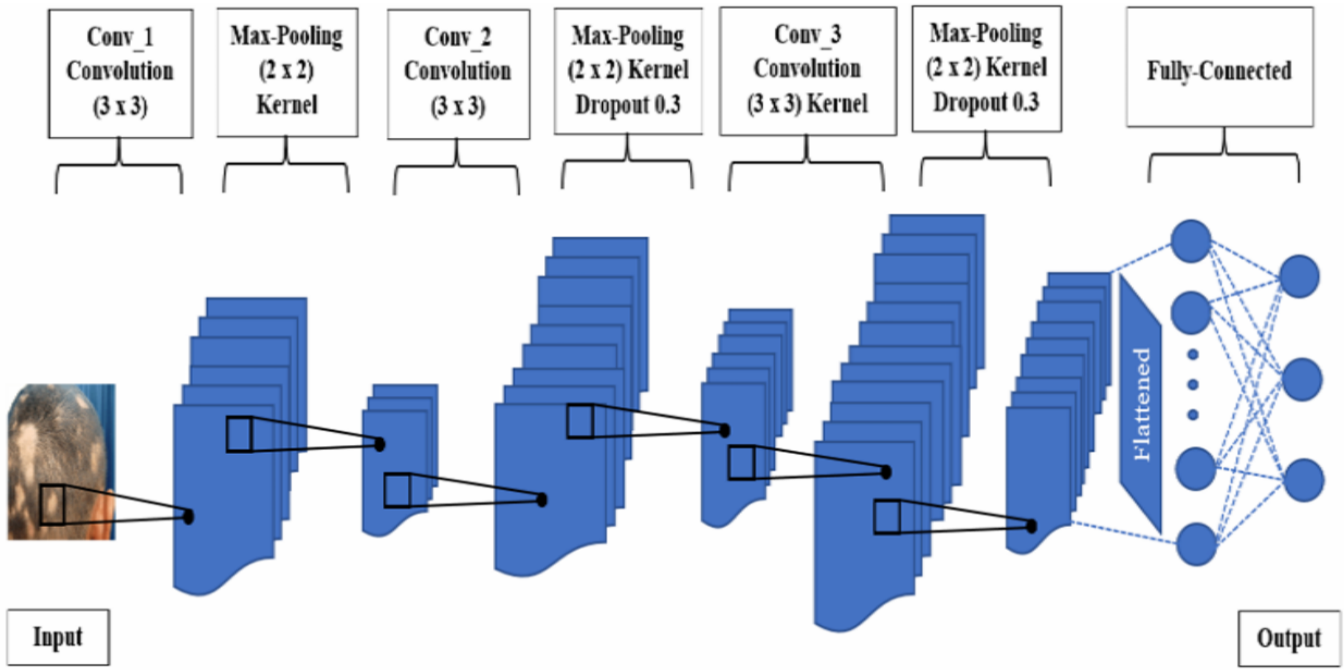


**Fig. 5:** Architecture of the Neural Network Model [13]

Our final model consists of three convolutional layers, each followed by a pooling layer to minimize overfitting and reduce feature sizes. A 3×3 kernel was applied across layers, and ReLU activation was used to introduce non-linearity while preventing computational inefficiencies. A dropout rate of o.3 was set to enhance generalization by randomly deactivating 30% of neurons during each epoch. The extracted feature maps were flattened into a one-dimensional vector before passing through two fully connected layers. The final classification was achieved using a softmax activation function, mapping outputs to three probability scores corresponding to alopecia, folliculitis, and psoriasis. We leveraged the Adam optimizer, known for its adaptive learning capabilities, to fine-tune weights and accelerate convergence. The dataset was split 70-30 into training and validation sets for training. The model was trained utilizing a batch size of 16 over 50 epochs, ensuring optimal performance and robustness in hair disease classification.

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

## IV. RESULTS AND DISCUSSIONS

This section will briefly discuss the CNN model's performance, which was trained using optimal hyperparameters determined through grid search. The dataset contained 105 randomly chosen photos for training and 45 for testing, which was divided into 70% training and 30% testing. After pre-processing, the model was trained on the dataset and evaluated on unseen test images. The system performed well, achieving 97% training and 92% validation accuracy. As illustrated in Fig. 6, the training loss reduced from 1.1685 to 0.1017, while the validation loss decreased from 1.1260 to 0.3438 over 50 epochs. Simultaneously, training accuracy improved from epoch 1 to epoch 50, reaching 97%, while validation accuracy rose to 92% which is shown in Fig. 7.

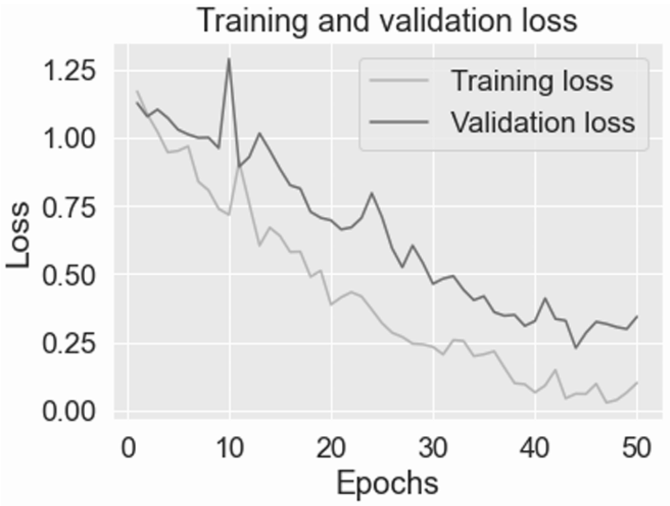


**Fig. 6:** Training and Validation loss for CNN model

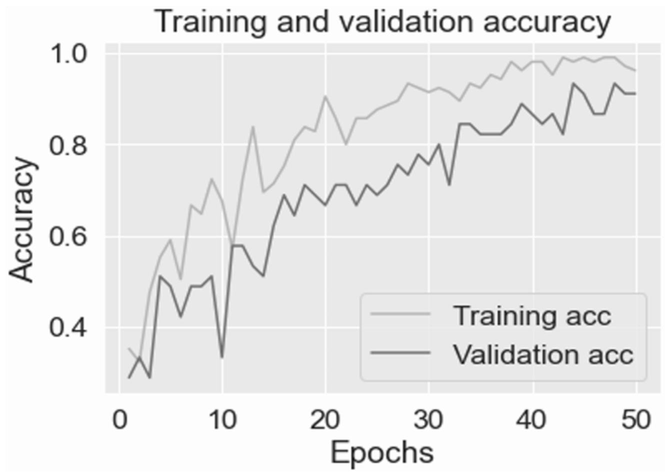


**Fig. 7:** Training and Validation Accuracy

The confusion matrix demonstrated in Fig. 8 provides a breakdown of classification outcomes. Among the 45 test images, the model correctly classified 17 out of 19 alopecia images, with two misclassified as psoriasis. For 13 psoriasis cases, 11 were correctly identified, while two were misclassified as alopecia. Notably, all 13 folliculitis images were classified accurately. Fig. 9 presents the fractional misclassification rate for each category.

Precision, recall, and F1-score metrics further quantify the model's performance, which is displayed in Fig. 10. The performance metrics for the model assessing skin diseases, specifically alopecia, psoriasis, and folliculitis, reveal important insights into its diagnostic capabilities. Alopecia shows impressive results with a precision of 0.96 and a recall of 0.97, leading to an F-score and accuracy of 0.97. These results demonstrate how well the program detects actual alopecia instances while reducing false positives. Since the model can detect almost every case of the illness, which is essential for timely treatment, the high recall is very noteworthy. Performance measures for psoriasis are somewhat lower, with precision at 0.90 and recall at 0.93, leading to an F-score of 0.91 and accuracy of 0.92. Although these results are commendable, they suggest potential progress in accurately diagnosing all cases of psoriasis. Poorer accuracy increases the likelihood of false positives, which can result in unnecessary

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

patient suffering. The accuracy and recall of psoriasis must be improved to improve the diagnostic reliability of the model.

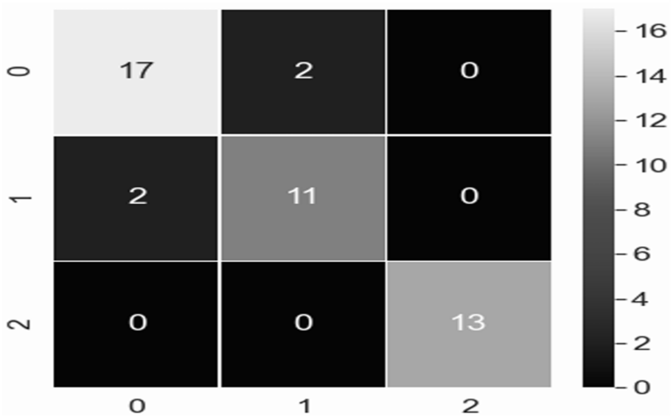


Fig. 8: Confusion Metrix for CNN model

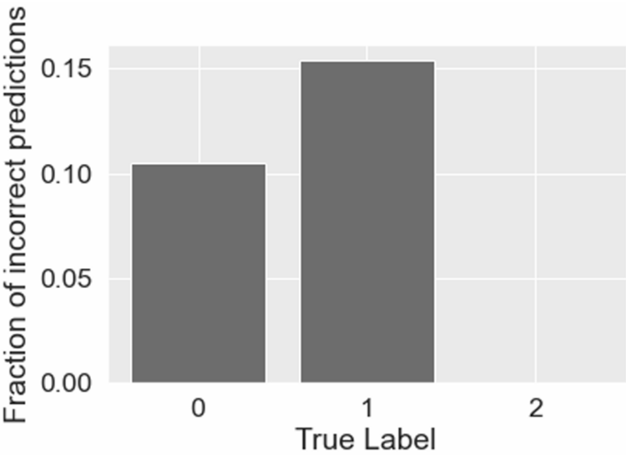


Fig. 9: Fractional Incorrect Prediction of CNN Model

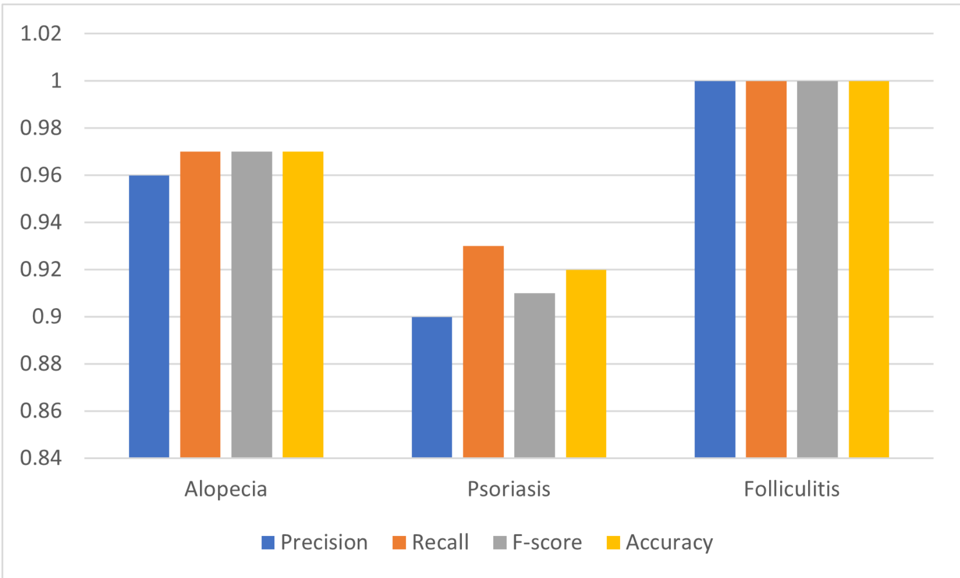


**Fig. 10:** Performance of the Model on Skin Diseases

Folliculitis works notably well, reaching 100% with accuracy, F- score, recall, and precision. By accurately identifying folliculitis, the model ensures that patients receive the correct medication without performing the danger of a false positive. This suggests that it is easier to differentiate symptoms of folliculitis from other skin disorders. Lastly, the model suggests areas for improvement in psoriasis detection while demonstrating strong diagnostic capabilities for skin conditions, especially folliculitis. Future studies should investigate cutting-edge machine-learning methods to improve prediction accuracy and broaden the training dataset to include a range of clinical presentations. Excellent performance in every scenario will guarantee automated diagnostic tools are trusted in clinical practice and enhance patient outcomes.

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025

## V. CONCLUSION AND FUTURE WORK

Hair loss and scalp disorders can frequently go undiagnosed because of a lack of knowledge and a drawn-out diagnostic process, even though early detection of hair and scalp-related disorders is essential to the treatment process. Early illness identification might be made easier with the help of an AI-based program. In this study, we fed 150 preprocessed picture data into a 2-D convolutional neural network model that reliably identifies three disorders associated with the hair and scalp: psoriasis, folliculitis, and alopecia. The remaining 30% of the photos were examined for model testing after 70% of the data had been used to train the model. This study demonstrates how well CNNs work for automatically identifying conditions affecting the hair and scalp, including folliculitis, psoriasis, and alopecia. Our model's remarkable training and validation accuracies of 97% and 92%, respectively, were attained by meticulously selecting a dataset of 150 photos and applying strong image preprocessing techniques. Our study establishes the foundation for incorporating AI into dermatology through more effective and precise disease detection, improving clinical decision-making and patient outcomes.

## REFERENCES

1. Banerjee, D., Kukreja, V., Bordoloi, D., & Choudhary, A. (2024). Enhanced hair disease classification using deep learning. *In 2024 11th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 58–63). IEEE.
2. Polak-Witka, K., Rudnicka, L., Blume-Peytavi, U., & Vogt, A. (2020). The role of the microbiome in scalp hair follicle biology and disease. *Experimental Dermatology, 29*(3), 286–294.
3. Dinnes, J., Deeks, J. J., Chuchu, N., di Ruffano, L. F., Matin, R. N., Thomson, D. R., Wong, K. Y., Aldridge, R. B., Abbott, R., & Fawzy, M. (2018). Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults. *Cochrane Database of Systematic Reviews, 12*.
4. Zhu, H., Isikman, S. O., Mudanyali, O., Greenbaum, A., & Ozcan, A. (2013). Optical imaging techniques for point-of-care diagnostics. *Lab on a Chip, 13*(1), 51–67.
5. Ahmad, Z., Rahim, S., Zubair, M., & Abdul-Ghafar, J. (2021). Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: Present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. *Diagnostic Pathology, 16*, 1–16.
6. Iqbal, M. I., & Kaushik, P. (2025). Deep learning for disease detection—A deep dive into deep learning techniques such as convolutional neural networks (CNNs) and their use in disease detection. *AI in Disease Detection: Advancements and Applications*, 99–122.
7. Lee, A. K. W., Chan, L. K. W., Lee, C. H., Bohórquez, J. M. C., Haykal, D., Wan, J., & Yi, K.-H. (2025). Artificial intelligence application in diagnosing, classifying, localizing, detecting and estimating the severity of skin conditions in aesthetic medicine: A review. *Dermatological Reviews, 6*(1), e70015.
8. Chowdhury, M. S., Sultan, T., Jahan, N., Mridha, M. F., Safran, M., Alfarhood, S., & Che, D. (2024). Leveraging deep neural networks to uncover unprecedented levels of precision in the diagnosis of hair and scalp disorders. *Skin Research and Technology, 30*(4), e13660.
9. Shakeel, C. S., Khan, S. J., Chaudhry, B., Aijaz, S. F., & Hassan, U. (2021). Classification framework for healthy hairs and alopecia areata: A machine learning (ML) approach. *Computational and Mathematical Methods in Medicine, 2021*(1), 1102083.
10. Kapoor, I., & Mishra, A. (2018). Automated classification method for early diagnosis of alopecia using machine learning. *Procedia Computer Science, 132*, 437–443.
11. ALEnezi, N. S. A. (2019). A method of skin disease detection using image processing and machine learning. *Procedia Computer Science, 163*, 85–92.
12. Hameed, N., Shabut, A. M., & Hossain, M. A. (2018). Multi-class skin diseases classification using deep convolutional neural network and support vector machine. *In 2018 12th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)* (pp. 1–7). IEEE.

Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03, Issue. 01, 2025
13. Roy, M., & Protity, A. T. (2022). Hair and scalp disease detection using machine learning and image processing. *arXiv preprint arXiv:2301.00122*.
14. Debbarma, T., Pal, T., Saha, A., & Debbarma, N. (2025). HCNNet: A hybrid convolutional neural network for abnormal human driver behavior detection. *Sādhanā, 50*(1), 9.
15. Gür, Y. E., Toğaçar, M., & Solak, B. (2025). Integration of CNN models and machine learning methods in credit score classification: 2D image transformation and feature extraction. *Computational Economics*, 1–45.
16. Madapuri, R. K., & Senthil Mahesh, P. C. (2017). HBS-CRA: Scaling impact of change request towards fault proneness: Defining a heuristic and biases scale (HBS) of change request artifacts (CRA). *Cluster Computing, 22*(S5), 11591–11599. https://doi.org/10.1007/s10586-017-1424-0
17. Dwaram, J. R., & Madapuri, R. K. (2022). Crop yield forecasting by long short-term memory network with Adam optimizer and Huber loss function in Andhra Pradesh, India. *Concurrency and Computation: Practice and Experience, 34*(27). https://doi.org/10.1002/cpe.7310
18. Thulasi, M. S., Sowjanya, B., Sreenivasulu, K., & Kumar, M. R. (2022). Knowledge attitude and practices of dental students and dental practitioners towards artificial intelligence. *International Journal of Intelligent Systems and Applications in Engineering, 10*(1s), 248–253.