IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14621

Deep Learning-Based Image Classification System for Scalp Diseases and Hair Loss Stages

Swarnalatha G L[1], Karuna M[2], Jeevitha S[3], Varshitha M V[4]

Assistant Professor, Dept. of CSE, Sri Siddhartha Institute of Technology, Tumakuru, Karnataka, India[1]

UG student, Dept. of CSE, Sri Siddhartha Institute of Technology, Tumakuru, Karnataka, India[2]

UG student, Dept. of CSE, Sri Siddhartha Institute of Technology, Tumakuru, Karnataka, India[3]

UG student, Dept. of CSE, Sri Siddhartha Institute of Technology, Tumakuru, Karnataka, India[4]

Abstract: This study presents a deep learning-based approach for automatic classification of scalp diseases and hair loss stages using image data. Leveraging convolutional neural networks (CNNs) with transfer learning, we evaluated multiple pre-trained models including ResNet50, VGG16, VGG19, and EfficientNet. Our method addresses challenges related to limited dataset size through image preprocessing and augmentation techniques, achieving high accuracy in distinguishing conditions like alopecia, psoriasis, and folliculitis, as well as hair loss progression stages. The trained models were integrated into a web application for user-friendly scalp condition diagnosis, enabling early detection and ongoing health monitoring.

Keywords: Deep Learning, Scalp disease, Hair Loss, Convolutional neural network (CNN).

I. INTRODUCTION

Hair and scalp health play an important role in a person's overall well-being and self-confidence. Conditions such as alopecia, psoriasis, folliculitis, and progressive hair loss are increasingly common among individuals of all ages. If left untreated, these issues can lead to serious complications like permanent hair loss, infections, or a negative impact on mental health. Diagnosing these conditions usually requires visiting a dermatologist, which may not be affordable or easily accessible for everyone.

To address this challenge, this research proposes a deep learning-based solution that can automatically detect scalp conditions and hair loss stages using image classification techniques. We used a convolutional neural network (CNN) architecture combined with transfer learning to improve efficiency, especially since our dataset is relatively small. Transfer learning allows the model to benefit from features learned by large pre-trained networks, reducing training time and improving accuracy. We tested multiple pre-trained models and selected the one that gave the best performance, achieving training accuracy of 98%, validation accuracy of 72%, F1-score of 0.65, precision of 0.73, recall of 0.62 and a loss value of 0.613.

The final model is deployed through a user-friendly web application that includes user authentication, image upload, and condition prediction functionalities. Users can register, log in, upload their scalp images, and instantly view the predicted scalp condition (alopecia, psoriasis, folliculitis) and hair loss stage (from 0 to 4, where stage 4 represents complete baldness). The application also includes a feature to track user history, helping individuals monitor improvements or worsening of their condition over time, supporting early intervention and consistent care.

II. RELATED WORK

[1] The research focuses on using deep learning techniques to classify hair loss levels based on facial images. The primary algorithm employed is a convolutional neural network (CNN),The dataset was derived from the CelebA database, which were manually annotated according to the Hamilton-Norwood scale. To enhance training, data augmentation techniques, such as horizontal flips, Gaussian noise, and contrast enhancement, were ap plied, resulting in multiple augmented datasets. The objectives include developing an automated, vision-based system for early detection and classification of hair loss levels, aiding medical, commercial, and security applications. However, drawbacks include limitations in dataset size, biases due to image occlusions, and lower accuracy for certain classes, particularly those representing early stages of hair loss. The study emphasizes the need for further data refinement to improve classification performance.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14621

[2] The paper presents a method for diagnosing scalp conditions using machine learning techniques. The researchers developed a system to classify conditions such as dandruff, alopecia, psoriasis, and oily scalp. They employed convolutional neural networks (CNNs) as the primary algorithm due to its strong performance in image classification tasks. The dataset used consists of scalp images collected from clinical sources and public repositories. The images were pre-processed and augmented to address challenges like class imbalance and overfitting. The primary objectives were to enable early diagnosis of scalp issues, create an efficient and user-friendly system for image-based detection, and improve classification accuracy compared to existing methods. However, the system faces draw backs such as limited generalizability due to the dataset's constraints, potential biases in data collection, and the need for more extensive real-world validation.

[3] The research paper explores the application of deep learning techniques in medical image analysis, specifically focusing on dermatological disease classification. It highlights the effectiveness of convolutional neural networks (CNNs) in diagnosing various skin and scalp conditions by analysing medical images. The study discusses the advantages of deep learning models such as VGGNet, ResNet, and MobileNet in accurately classifying dermatological conditions, emphasizing their high performance in pattern recognition and feature extraction. The paper also examines the importance of preprocessing techniques, including image enhancement, segmentation, and noise reduction, in improving classification accuracy. Additionally, it emphasizes the role of dataset augmentation methods, such as rotation and flipping, to enhance model generalization. Another critical aspect covered is the necessity of explainability in AI-driven medical diagnostics, with techniques like Grad CAM being suggested for model interpretability.

[4] The research paper explores advanced machine learning techniques for medical image analysis, focusing on the classification of scalp conditions. The study highlights the significance of early and accurate diagnosis of dermatological issues to improve patient outcomes. The researchers employ deep learning models, particularly Convolutional Neural Networks (CNNs), which have demonstrated superior performance in image classification tasks. The study utilizes datasets containing scalp images labelled with conditions like dandruff, psoriasis, and alopecia, among others. The authors experiment with multiple architectures, including ResNet, VGG, and EfficientNet, to compare their effectiveness in feature extraction and classification. They also apply preprocessing techniques such as image augmentation, noise reduction, and contrast enhancement to improve model performance. Additionally, the paper discusses the role of transfer learning in improving accuracy when dealing with limited datasets. Evaluation metrics like accuracy, precision, recall, and F1-score are used to assess model performance.

III. METHODOLOGY

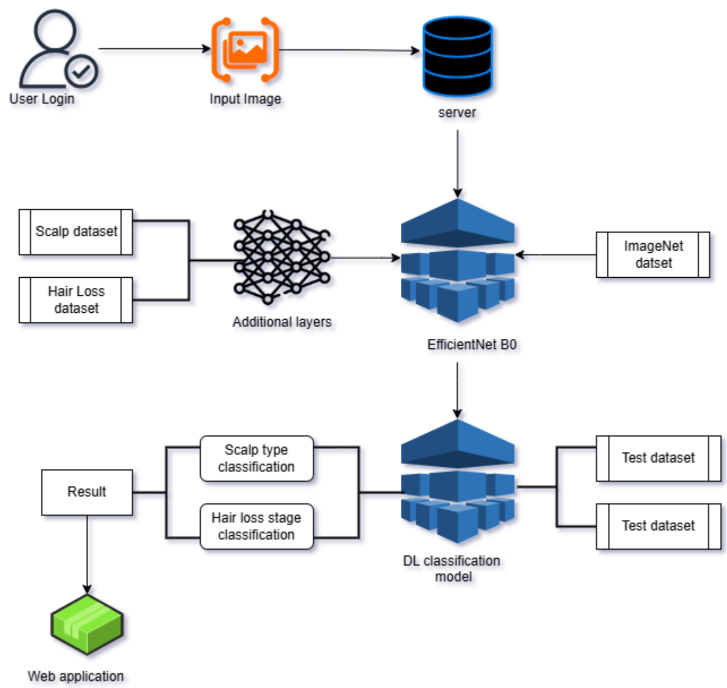

Fig. 1 Architecture of the proposed system

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14621

1. Data collection and preparation:

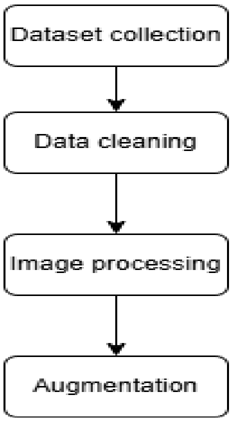

Fig. 2 Flow chart of Data collection and preparation process

a) Dataset collection: Two distinct datasets are curated, one focusing on various stages of hair loss and the other on different scalp conditions such as psoriasis, alopecia and folliculitis. The entire dataset is collected using custom methods, including manual sourcing of scalp images from open-access platforms and dermatology-related content shared under fair use**.** 50 images are provided for each class

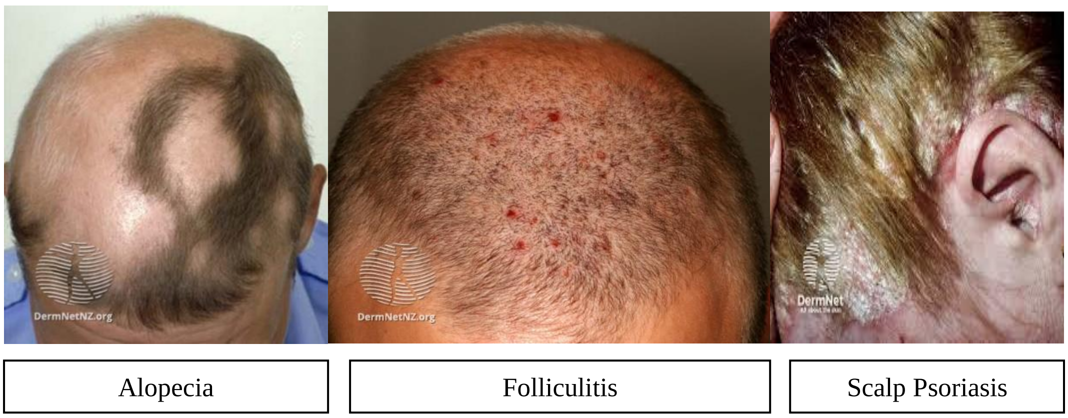

Fig.3 A sample scalp dataset

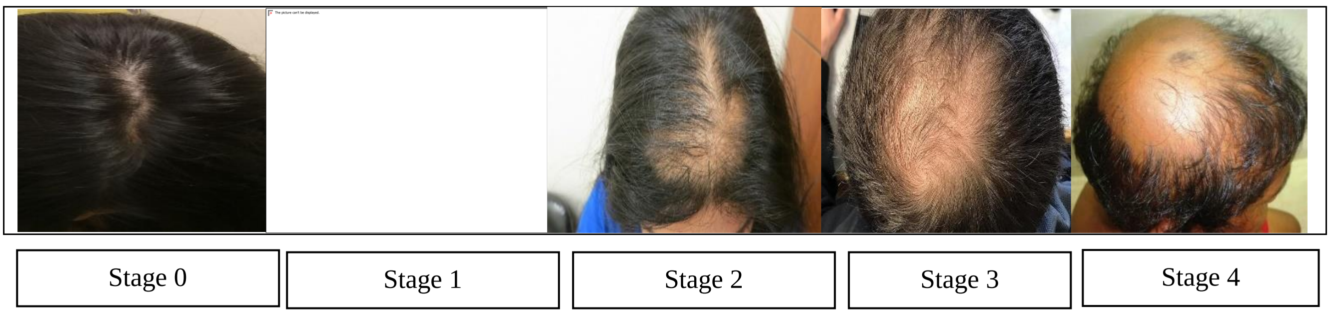

Fig.4 A sample Hair Loss dataset

b) **Data cleaning:** All images undergo careful inspection to eliminate corrupted, low-quality, or duplicate files that may hinder model learning. Labels assigned to each image are manually verified to ensure accuracy and consistency, maintaining the integrity of the classification process.

c) **Image Preprocessing**: Each image is resized to a fixed dimension (typically 224x224 pixels) suitable for input into a convolutional neural network (CNN).Pixel values are normalized to a standard scale (0 to 1) to improve training stability and model performance. The input image is then converted to HSV to help the model better detect visual cues like colour changes or patches, reduce the effect of lighting differences, and ultimately improve classification accuracy.

d) Data Augmentation: To simulate real-world image variability and improve model generalization, several augmentation techniques are applied, such as rotation, flipping, zooming, and brightness adjustments. This process also helps in balancing class distributions, especially for under-represented scalp conditions, making the model more robust and reliable.

2. Model Selection and Training

Fig. 5 Flowchart of model selection and training process

a) Model selection: A variety of powerful and widely used pre-trained convolutional neural network (CNN) architectures are selected for the classification tasks. These include ResNet50, VGG16, VGG19, and EfficientNet. Each of these models is known for its unique structure and performance on image classification benchmarks. The method of transfer learning is employed, where the base layers of these pre-trained models originally trained on large-scale datasets like ImageNet are retained.2 dense hidden layers and 1 dense output layer with 128 fully connected neurons are added while retraining the model with custom scalp and hair loss dataset. This approach significantly reduces the need for a large dataset and shortens the training time while improving performance**.**

b) **Training Process:** Each selected CNN model is fine-tuned separately on both the hair loss dataset and the scalp condition dataset. This ensures the models learn patterns specific to each task without interference. The training process uses categorical cross-entropy as the loss function, which is suitable for multi-class classification problems where each input belongs to one of several distinct categories. The entire dataset is split into three subsets -70 percent for training, 20 percent for validation, and 10 percent for testing. This split allows effective model training, parameter tuning through validation, and final evaluation on unseen data.

c) **Evaluation Metrics**: To assess the performance of each model, several metrics are considered: accuracy, precision, recall, F1-score, and loss curves. These metrics provide a comprehensive evaluation of the model's effectiveness across different classes.. Based on the results observed during validation, the best-performing models are selected. This helps in ensuring the deployed model delivers accurate and reliable predictions across all scalp and hair loss categories.

d) **Optimization**: To prevent the model from overfitting the training data and to improve generalization to new inputs, techniques such as dropout and early stopping are used. Further hyper-parameter tuning is carried out by experimenting with various values of learning rate, batch size, and the number of training epochs. This helps in finding the best configuration for stable and high-performing models.

3. System Design: The system design consists of both frontend and backend components. The frontend was developed using HTML, CSS, and JavaScript to create an intuitive user interface that allows users to easily navigate the application and upload scalp images. On the backend, a Flask framework was used to manage image uploads, perform model inference for predicting scalp conditions and hair loss stages, and display the results along with relevant recommendations. The trained deep learning models were integrated into the backend to ensure smooth and accurate analysis. Finally, the complete system was deployed on a local server for testing purposes.

-

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 6, June 2025
DOI: 10.17148/IJARCCE.2025.14621

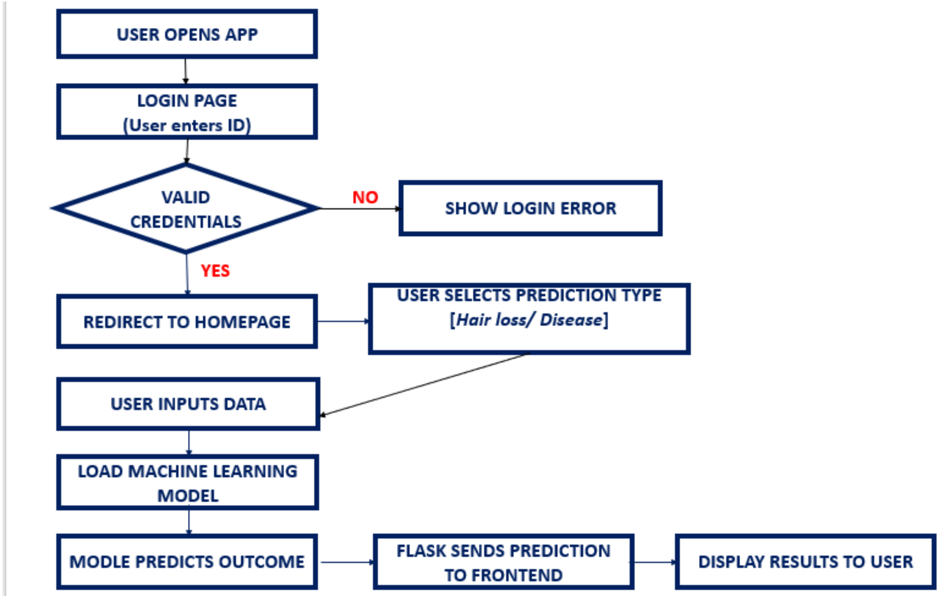

Fig.6 Flowchart of the System design

a) **User opens the application**: The system is accessible via a web browser, allowing users to start the prediction process by navigating to the application's homepage.

b) **Login Page**: Upon opening the application, users are directed to a login page where they must enter their credentials, such as a user ID and password.

c) **Credential Verification**: Once the credentials are entered, the front-end sends the login details to the backend for validation. If the entered details are incorrect, an error message is displayed, prompting the user to re-enter their credentials. If authentication is successful, the user is redirected to the homepage, gaining access to the system's features.

d) **Homepage & Prediction Type Selection**: After successful login, the user lands on the homepage, where they must choose the type of scalp health assessment they want to perform. The system provides options for diagnosing hair loss or identifying specific scalp conditions. The selection ensures that the Deep Learning model processes the image accordingly for precise results.

e) **Data Input Section**: Users are then prompted to upload an image of their scalp for analysis. A user-friendly interface allows easy file selection and submission, ensuring a smooth experience.

f) **Load Machine Learning model**: the user data is passed to Deep Learning model through flask for prediction.

g) **Model predicts outcome**: model predicts the scalp condition/hair loss stage using the user's input data.

h) **flask sends prediction to front end**: on results prediction, flask interacts with the Machine learning model and displays the results on the front-end to users.

IV. RESULTS AND DISCUSSION

Metric	Value
Training accuracy	98%
Validation accuracy	72%
Precision	0.73
Recall	0.62
F1-score	0.65
Loss	0.136

Fig.7 Evaluation metrics of EfficientNet model

DOI: 10.17148/IJARCCE.2025.14621

Snapshots of web application:

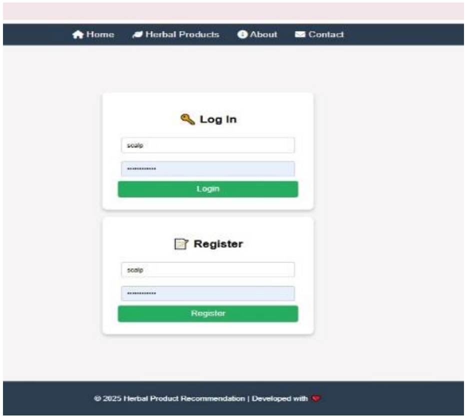

Fig.8 Login page of web application

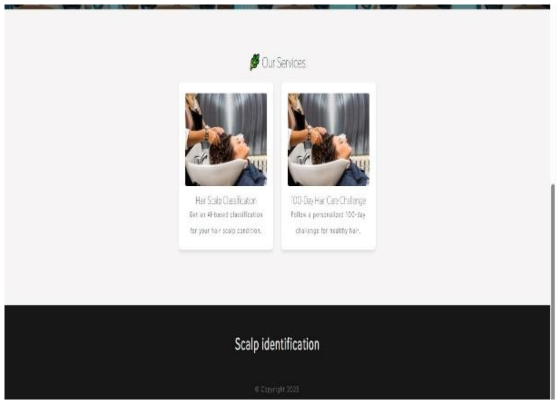

Fig. 9. Home page of web application.

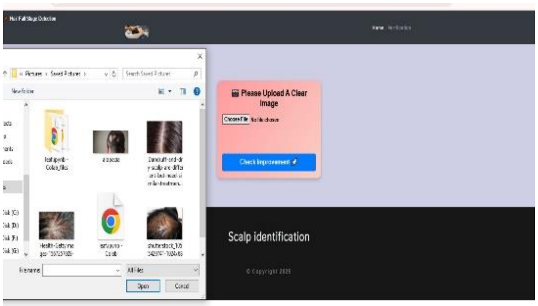

Fig.10. Image upload page of web application

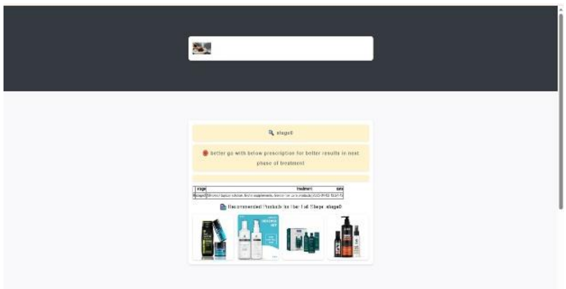

Fig.11. Result page of web application.

EfficientNet was selected as the model for the web application due to its superior training performance compared to the other models evaluated. However, refering to Fig 4. presenting its evaluation metrics, the gap between training and validation results indicates potential overfitting. This suggests that while EfficientNet learns the training data well, its ability to generalize to unseen data is limited.

To address these issues, future work will focus on increasing the size and diversity of the training dataset, applying regularization methods such as dropout or L2 regularization, and exploring simpler model architectures. Additionally, techniques like data augmentation or resampling to balance classes could improve recall and overall model robustness.

V. CONCLUSION

In this study, a Convolutional Neural Network (CNN) architecture integrated with transfer learning was utilized to address a classification problem effectively. The objective was to build a model capable of learning complex features from the input data while reducing training time and improving generalization using pre-trained weights.

The model achieved a validation accuracy of 72%, with a precision of 0.73, recall of 0.62, and an F1-score of 0.65. These results indicate that the model performs moderately well, particularly in minimizing false positives. However, the recall score suggests that the model misses some true positive cases, which could be critical depending on the application.

Despite this limitation, the use of transfer learning significantly contributed to feature extraction and overall performance, especially in a scenario with limited data. The findings suggest that CNNs, combined with transfer learning, offer a promising direction for classification tasks. Future improvements may include dataset balancing, hyperparameter tuning, or experimenting with deeper architectures to further boost recall and overall model robustness.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14621

REFERENCES

[1] H. Benhabiles, K. Hammoudi, Z. Yang, F. Windal, M. Melkemi, F. Dornaika, and I. Arganda-Carreras, "Deep Learning based Detection of Hair Loss Levels from Facial Images," in *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Istanbul, Turkey, Nov. 2019, pp. 1–6, DOI: 10.1109/IPTA.2019.8936122.

[2] W.-J. Chang, L.-B. Chen, M.-C. Chen, Y.-C. Chiu, and J.-Y. Lin, "ScalpEye: A Deep Learning-Based Scalp Hair Inspection and Diagnosis System for Scalp Health," *IEEE Access*, vol. 8, pp. 134826–134837, 2020, DOI: 10.1109/ACCESS.2020.3010847.

[3] R. P. Suryavanshi, A. Wanjari, M. Mahajan, T. Kachewar, and V. Khobragade, "Hair and Scalp Disease Detection using Machine Learning and Image Processing," *International Journal of All Research Education and Scientific Methods (IJARESM)*, vol. 11, no. 9, pp. 1701, Sep. 2023. Available: www.ijaresm.com.

[4] K. Sultanpure, B. Shirsath, B. Bhande, H. Sawai, S. Gawade, and S. Samgir, "Hair and Scalp Disease Detection Using Deep Learning," *arXiv preprint arXiv:2403.07940*, Mar. 2024.

[5] I. Kapoor and A. Mishra, "Automated Classification Method for Early Diagnosis of Alopecia Using Machine Learning," in *Proc. International Conference on Computational Intelligence and Data Science (ICCIDS)*, 2018, Amity University, Noida, India.

[6] B. S. H. Reddy, "Deep Learning-Based Detection of Hair and Scalp Diseases Using CNN and Image Processing," *Milestone Research Publications*, DOI: 10.5281/zenodo.14965660, received Feb. 13, 2025, revised Feb. 28, 2025, accepted Mar. 4, 2025.

[7] A. K. W. Lee, L. K. W. Chan, C. H. Lee, J. M. Cerón Bohórquez, D. Haykal, J. Wan, and K.-H. Yi, "Artificial Intelligence Application in Diagnosing, Classifying, Localizing, Detecting and Estimating the Severity of Skin Conditions in Aesthetic Medicine: A Review," *Dermatological Reviews*, vol. 2025, Article ID 70015, Jan. 2025, DOI: 10.1002/der2.70015.