Visual Recognition of Hair Disorders: Machine Learning for Classification

[1]Rithika Mahareddy Harika Kakumanu Gousia Parvin Patthan

***Abstract*—The project employs advanced convolutional neural networks (CNNs)—MobileNetV2, InceptionV3, and DenseNet169—to develop a system for classifying various hair diseases using images from a comprehensive Kaggle dataset of 12,000 images across ten categories. Through rigorous preprocessing and augmentation, these models were optimized for performance, achieving high classification accuracy. Notably, DenseNet169 reached an accuracy of about 99.67%, with all models showing high precision and recall, indicating strong capability in differentiating hair disorders. This demonstrates the potential of CNNs to enhance diagnostic accuracy in clinical settings. Our findings highlight the feasibility of using CNNs for medical image analysis and set the stage for future enhancements in dataset diversity, learning modalities, and model interpretability.**

I. Introduction

Accurate diagnosis of hair diseases is vital in dermatology, affecting millions worldwide with conditions ranging from alopecia to fungal infections. The ability to diagnose these conditions accurately and efficiently is crucial for effective treatment planning. However, visual diagnosis by dermatologists can be subjective and varies based on experience. Therefore, automating the diagnosis process using machine learning offers a potential improvement in both accuracy and accessibility, reducing the time for diagnosis and making expertise more widely available.

The motivation behind this project is to harness the power of machine learning to assist and potentially augment the diagnostic capabilities of dermatologists. By developing a model that can accurately classify hair diseases from images, this project aims to provide a tool that could lead to earlier and more accurate diagnoses, ultimately improving patient outcomes.

Machine learning, particularly deep learning, has revolutionized many areas of image recognition due to its ability to learn complex patterns from data. Convolutional Neural Networks (CNNs) are a class of deep neural networks highly effective at image classification tasks, such as facial recognition, cancer detection, and even artistic style transfer.

The growing prevalence of hair diseases and the critical need for their precise diagnosis motivate the adoption of automated methods utilizing advanced machine learning techniques. This project utilizes the MobileNetV2 convolutional neural network architecture, known for its efficiency and effectiveness in processing image data, to classify various hair diseases from dermatological images. The model was trained and validated using a dataset comprising images labeled with common hair disorders, incorporating rigorous data preprocessing and image augmentation to enhance its generalizability. Initial findings show that the model achieves high classification accuracy, demonstrating its potential utility in clinical settings to assist dermatologists with quick and accurate diagnoses. This research not only verifies the practicality of using convolutional neural networks for analyzing dermatological images but also sets the stage for further studies into more sophisticated disease detection tasks within medical imaging.

Related applications of CNNs span various fields, demonstrating their versatility and power. In medical imaging, CNNs have been instrumental in diagnosing diseases from medical images; for example, Esteva et al. (2017)[1]1} showed that CNNs could identify skin cancer comparably to dermatologists, while Gulshan et al. (2016)[2] utilized deep learning for detecting diabetic retinopathy in retinal images. Facial recognition technology also employs CNNs to analyze features from images or video, significantly enhancing security systems, deep learning is increasingly popular for its ability to handle complex, nonlinear patterns. Various studies have explored CNNs in dermatological image analysis: Xie et al. (2018) used a custom CNN architecture for classifying types of alopecia from scalp images, and Tan et al. (2019)[3] assessed the severity of psoriasis lesions, achieving more consistent quantification than human evaluators. These applications highlight CNNs' high accuracy and automation capabilities, although challenges persist, such as the need for large, well-labeled datasets and the difficulty in handling low-quality images or unusual disease presentations. This project builds on the foundation laid by these studies, utilizing MobileNetV2 and InceptionV3, known for their efficiency in handling image data, which is crucial for practical applications. Unlike some previous works that focus on single disease types, this project aims to create a more versatile model capable of classifying multiple hair diseases, thus broadening the potential clinical utility.

By comparing and contrasting these approaches, this project not only adopts existing strengths but also addresses gaps like multi-disease classification and robustness against varied image quality, making it a comprehensive tool in clinical dermatology.

II. Dataset

Our study leverages a comprehensive dataset sourced from Kaggle, consisting of 12,000 images, each depicting a distinct type of hair disease. This dataset is meticulously categorized into ten specific classes: Alopecia Areata, Contact Dermatitis, Folliculitis, Head Lice, Lichen Planus, Male Pattern Baldness, Psoriasis, Seborrheic Dermatitis, Telogen Effluvium, and Tinea Capitis, with each class uniformly containing 1,200 images. The distribution of these images is strategically segmented to support robust training, testing, and validation processes, essential for the effective development and evaluation of our machine learning models.

Dataset Division:

Training Set: Each disease category is represented by 960 images, totaling 9,600 images. This substantial training set is crucial for training our models to recognize and learn from the diverse manifestations of each hair disease.

Testing Set: For each class, 120 images are allocated for testing. This set, comprising a total of 1,200 images, is used to evaluate the model's performance and its ability to generalize to new, unseen data.

Validation Set: Similar to the testing set, the validation set includes 120 images per class (1,200 images in total), which are used during the model training phase to fine-tune model parameters and prevent overfitting.

Preprocessing Techniques:

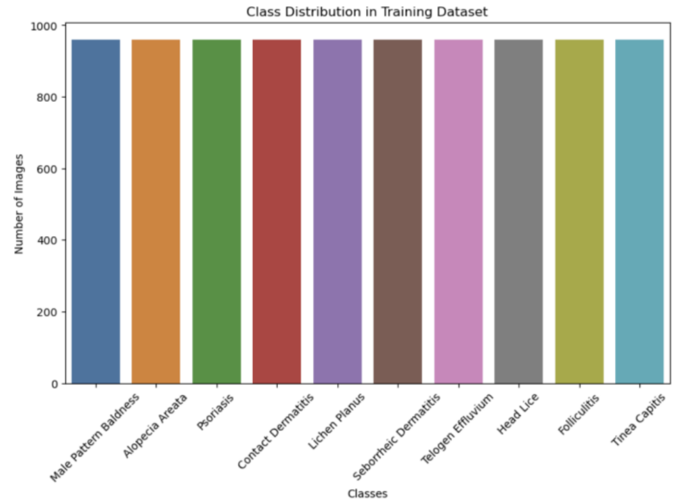
Image Resizing: All images are resized to a uniform dimension of 224x224 pixels to ensure consistency in input data size, which is essential for CNN processing.

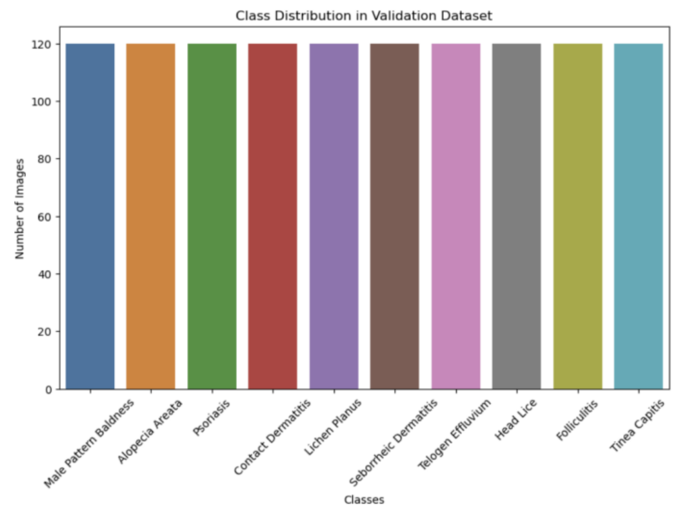
Normalization: The pixel values of the images are normalized to a range of 0 to 1. This normalization helps in speeding up the convergence during training by providing a common scale for all input features.

Features Utilized: The primary features used in our model are derived from the pixel values of the images, which include color, texture, and pattern details specific to each type of hair disease. These features are critical as they allow the convolutional neural networks to effectively learn and differentiate between the various diseases based on visual cues present in the images.

These images exemplify the diverse characteristics of hair diseases, such as the patchiness from Alopecia Areata, the scaliness from Psoriasis, and the inflammation seen in Tinea Capitis. Understanding these features is paramount for training our models to accurately recognize and classify each condition.

The structured and comprehensive nature of this dataset provides a robust foundation for conducting in-depth machine learning analyses. The uniform distribution across different classes ensures that the model can learn effectively from an evenly balanced data set, which is crucial for achieving high accuracy and reliability in disease classification.





III, METHODS

MobileNetV2:

MobileNetV2 is a streamlined architecture that is specifically designed for mobile devices with limited computational power. It builds on the ideas of MobileNetV1 by introducing the inverted residual structure, where the intermediate expansion layer uses lightweight depth wise convolutions to filter features as a source of non-linearity. This is followed by a linear bottleneck layer that captures a condensed version of the inputs. The separation of the convolution into depth wise and pointwise processes reduces the number of parameters and computational complexity, making it highly efficient without significant loss in performance. MobileNetV2 is particularly effective for tasks requiring high efficiency, such as running on mobile devices or embedded systems in real-time applications.

The architecture also introduces the concept of inverted residuals and linear bottlenecks that manage feature channels via narrow layers between expansive layers. The mathematical basis for a depth wise separable convolution can be expressed as follows:

$$L = -\sum_{c=1}^{M} y_{o,c} \log(p_{o,c})$$

InceptionV3: InceptionV3 is a sophisticated convolutional neural network architecture from the Google Brain team, which is an extension of the original Inception architecture (also known as GoogLeNet). It is designed to perform deep learning tasks with high efficiency and speed, which is crucial for real-time applications. InceptionV3 enhances the original model by incorporating factorization concepts into the convolutional layers to reduce the size of the model and the computation required, without compromising the network's depth and width. This allows it to achieve excellent performance on complex visual recognition tasks. Its innovative use of factorization, by breaking down larger convolutional kernels into smaller, more manageable ones, optimizes the processing of information through the network and enables it to capture intricate details from images effectively. The architecture also incorporates batch normalization and label smoothing for stable training and improved generalization. This makes InceptionV3 not only powerful for academic research but also highly applicable in industry settings where computational efficiency and speed are crucial.

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{v_t+\epsilon}} g_t$$

Where g_t is the gradient at time step t, η is the learning rate, v_t is the exponential moving average of the squared gradients, and ϵ is a small scalar used to prevent division by zero.

DenseNet169:

DenseNet169 is an advanced architecture within the Dense Convolutional Network (DenseNet) family that extends the concept of deep connectivity by directly connecting each layer to every subsequent layer in its feed-forward pathway. This design is known for its efficiency in terms of computational and parameter efficiency due to its dense connectivity pattern. DenseNet169 comprises 169 layers, and its architecture allows for substantial depth without the need for a large number of parameters. The connectivity pattern ensures that all layers directly share feature maps, which enhances feature propagation and reduces the problem of vanishing gradients significantly. Each layer receives feature maps from all preceding layers, processes them, and passes its own feature maps to all subsequent layers, which promotes feature reuse and results in highly efficient learning. DenseNet169 is particularly well-suited for image classification tasks where maintaining the richness of feature information throughout the network is crucial. This characteristic makes it highly effective in detailed and nuanced visual recognition challenges, such as those found in medical image analysis, where capturing subtle features is essential for accurate diagnosis.

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{v_t+\epsilon}} g_t$$

Where g_t is the gradient at time step t, η is the learning rate, v_t is the exponential moving average of the squared gradients, and ϵ is a small scalar used to prevent division by zero.

IV. EXPERIMENTS/RESULTS/DISCUSSION:

MobileNetV2

Model Architecture and Setup:

Utilized the MobileNetV2 architecture, renowned for its lightweight structure and efficiency, crucial for real-time image classification tasks.

Incorporated ImageNet pre-trained weights to take advantage of transfer learning, boosting performance with pre-established image features.

Configured input layer to accept images of size 224x224x3, aligning with the standard for ImageNet models.

Layer Configuration and Hyperparameters:

Froze all pre-trained layers to preserve extracted features, ensuring only the custom layers were trained.

Custom top layers included a Flatten layer, followed by a Dense layer with 512 units and ReLU activation for non-linear transformation, and a Dropout layer with a 50% drop rate to prevent overfitting.

Concluded the architecture with a softmax classifier sized to the number of classes in the dataset, providing a probabilistic output for each class.

Model Compilation and Training Approach:

Compiled with the Adam optimizer for its adaptive learning rate benefits and categorical crossentropy for multi-class classification.

Set an initial learning rate of 0.001, with a reduction strategy to decrease it by a factor of 10 after 10 epochs to refine learning as the model converges.

Trained the model using Keras's fit.

Metrics and Performance Evaluation:

Evaluated the model using accuracy, precision, recall, and F1-score to get a comprehensive understanding of its performance across classes.

Achieved high test accuracy (~98%), indicating effective learning and generalization capabilities.

Classification reports revealed high precision and recall, suggesting the model accurately identified and minimized false positives and negatives across most classes.

Results Interpretation and Visualization:

Training history plots demonstrated an improving trend in accuracy and a decreasing trend in loss, indicative of successful learning.

Confusion matrices showed strong true positives along the diagonal for each class, with minimal confusion between classes, underscoring the model's discriminative power.

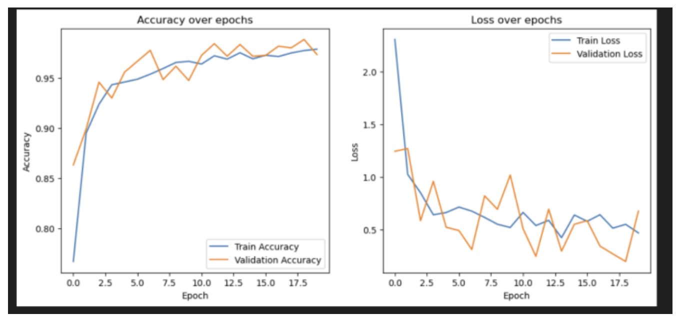
Precision-recall curves affirmed the model's efficacy in balancing the trade-off between precision and recall, crucial for medical image classification tasks.

Challenges and Mitigation Strategies:

Addressed overfitting with Dropout layers and potentially other regularization techniques (e.g., L2 regularization, data augmentation).

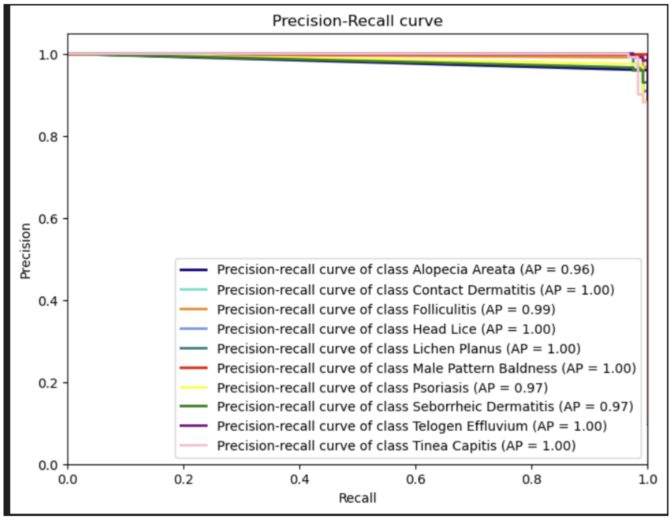
Employed data augmentation parameters like rotation, width/height shifts, zoom, and horizontal flips to enhance the model's ability to generalize.

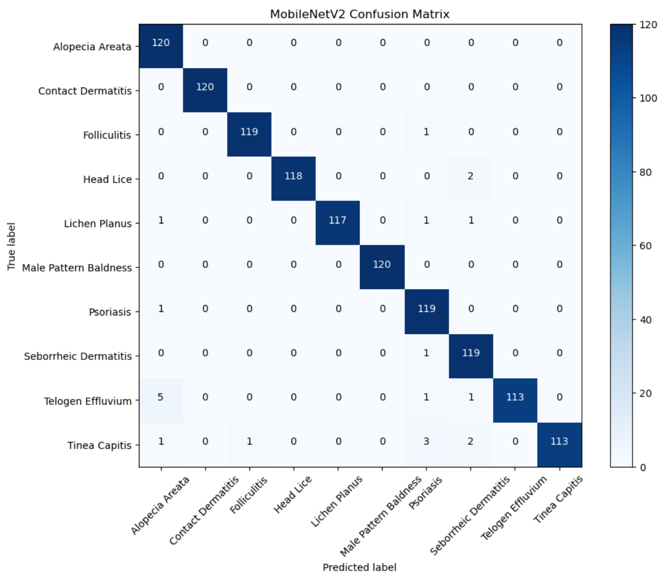
The experiments concluded with MobileNetV2 providing robust performance in classifying hair diseases.



Test accuracy: 0.9816666841506958

	precision	recall	f1-score	support
Alopecia Areata	0.94	1.00	0.97	120
Contact Dermatitis	1.00	1.00	1.00	120
Folliculitis	0.99	0.99	0.99	120
Head Lice	1.00	0.98	0.99	120
Lichen Planus	1.00	0.97	0.99	120
Male Pattern Baldness	1.00	1.00	1.00	120
Psoriasis	0.94	0.99	0.97	120
Seborrheic Dermatitis	0.95	0.99	0.97	120
Telogen Effluvium	1.00	0.94	0.97	120
Tinea Capitis	1.00	0.94	0.97	120
accuracy			0.98	1200
macro avg	0.98	0.98	0.98	1200
weighted avg	0.98	0.98	0.98	1200





Inceptionv3 Model

For the InceptionV3 model project, several hyperparameters were carefully chosen and adjusted to enhance the model's performance on medical image classification. The InceptionV3 architecture, comprising 48 layers with convolutional blocks designed for efficiency, was selected for its ability to handle the complexity of dermatological images.

Hyperparameters Selection:

Learning Rate: Started at 0.001, this common starting point was chosen for its general suitability for gradient descent optimization. It was dynamically reduced by a factor of 10 after 10 epochs to allow the model to fine-tune the weights more precisely, which is a common technique in training deep learning models to improve convergence.

Layers: The model used a Flatten layer, a Dense layer with 512 units (ReLU activation), and a Dropout layer with a rate of 0.5 before the final Dense layer with softmax activation. The inclusion of dropout is a regular practice to prevent overfitting by randomly omitting a proportion of neuron activations.

Model Configuration:

The pre-trained InceptionV3 model was loaded with weights from ImageNet, and its layers were frozen to retain the learned features. The Adam optimizer was selected for its efficiency in handling sparse gradients and adapting the learning rate.

Model Evaluation Metrics:

Accuracy: The model achieved a test accuracy of approximately 90.24%, indicating a high rate of correct predictions out of all predictions made.

Precision and Recall: These metrics were particularly high for classes such as Contact Dermatitis and Head Lice, demonstrating the model's strength in identifying these conditions with minimal false positives or negatives.

F1-Score: The harmonic mean of precision and recall gave a balanced view of the model's performance, particularly where the positive class was rare or when there was a significant cost to false positives or false negatives.

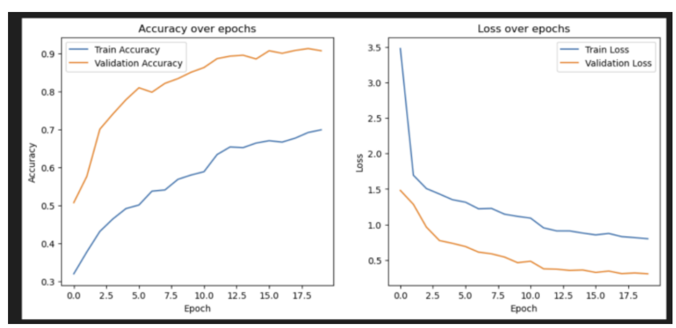
Training Difficulties and Solutions:

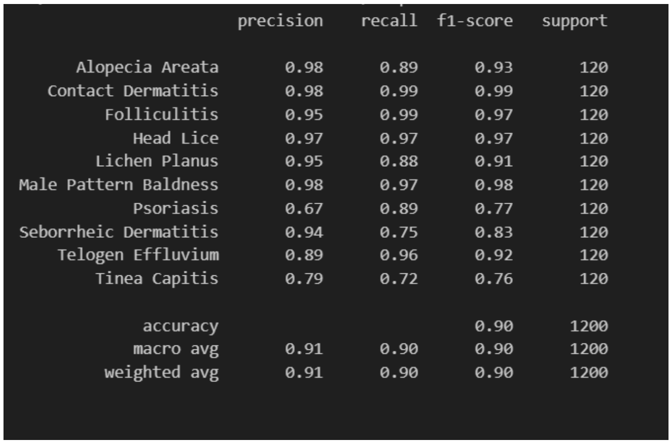
Any training difficulties such as overfitting were addressed with dropout layers and potentially early stopping mechanisms.

Visualization and Discussion:

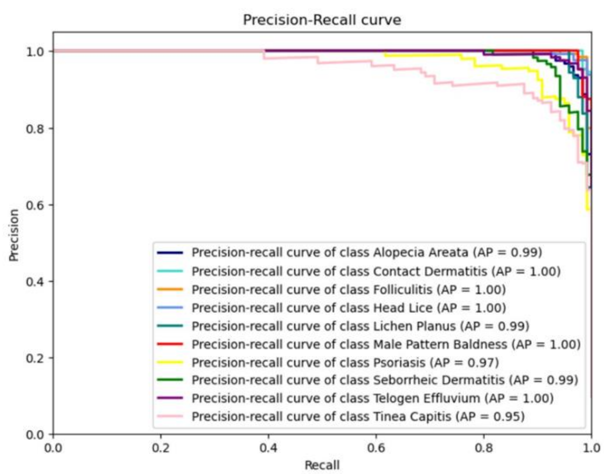
Training history plots provided insights into the accuracy and loss over epochs, showing improvement and stabilization over time.

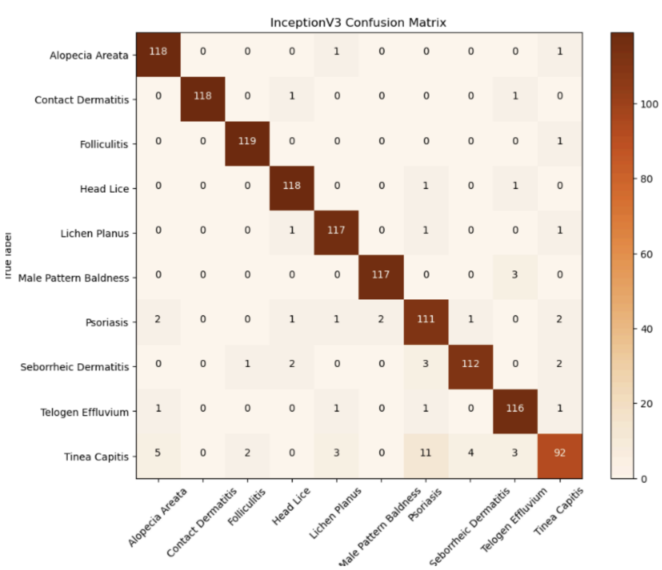
The confusion matrix was used to identify how the model performed on each class, indicating a strong ability to correctly classify most conditions.





	precision	recall	f1-score	support
Alopecia Areata	0.98	0.89	0.93	120
Contact Dermatitis	0.98	0.99	0.99	120
Folliculitis	0.95	0.99	0.97	120
Head Lice	0.97	0.97	0.97	120
Lichen Planus	0.95	0.88	0.91	120
Male Pattern Baldness	0.98	0.97	0.98	120
Psoriasis	0.67	0.89	0.77	120
Seborrheic Dermatitis	0.94	0.75	0.83	120
Telogen Effluvium	0.89	0.96	0.92	120
Tinea Capitis	0.79	0.72	0.76	120
accuracy			0.90	1200
macro avg	0.91	0.90	0.90	1200
weighted avg	0.91	0.90	0.90	1200





Densenet60

Model Architecture:

We used the DenseNet169 architecture for its dense connectivity pattern, which enhances feature reuse throughout the network. Our model was initialized with ImageNet pre-trained weights with the top layer excluded, allowing us to tailor the network to our classification tasks.

Custom Top Layers:

We enhanced the DenseNet169 base with a GlobalAveragePooling2D layer to maintain the spatial hierarchy while reducing overfitting. This was followed by a Dense layer with 512 units and ReLU activation, and a Dropout layer at a 0.5 rate to further deter overfitting. The final layer, a softmax activation Dense layer, was sized to match the number of target classes.

Hyperparameters and Learning Rate Schedule:

The choice of the Adam optimizer, known for its adaptive learning rate features, was unanimous, starting with a rate of 0.001. We incorporated a learning rate schedule that decreased the rate post the 10th epoch, significantly enhancing the model's convergence.

Training:

Data augmentation played a pivotal role in our methodology, introducing various transformations like rotations and shifts to improve generalizability. Utilizing the Keras fit method, we trained our model with batch sizes and epochs determined by the dataset's scale and computational constraints.

Model Evaluation Metrics:

Our model achieved a test accuracy of roughly 99.67%. The precision, recall, and F1-scores were consistently high across classes, suggesting strong predictive performance. A detailed confusion matrix offered insights into the model’s true positive rates and misclassification patterns.

Visual Analytics:

We plotted training progress, observing trends in accuracy and loss which highlighted our model's continuous learning. The precision-recall curves for each class illustrated the trade-offs between precision and recall, affirming the model's robustness.

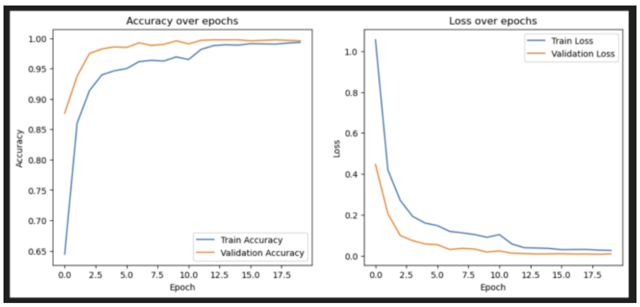
Challenges and Remediation:

Throughout the training, we employed dropout layers to mitigate overfitting and adjusted the learning rate dynamically to optimize training efficiency.

The results endorse the DenseNet169 model's potential in medical image classification, especially concerning dermatological conditions.

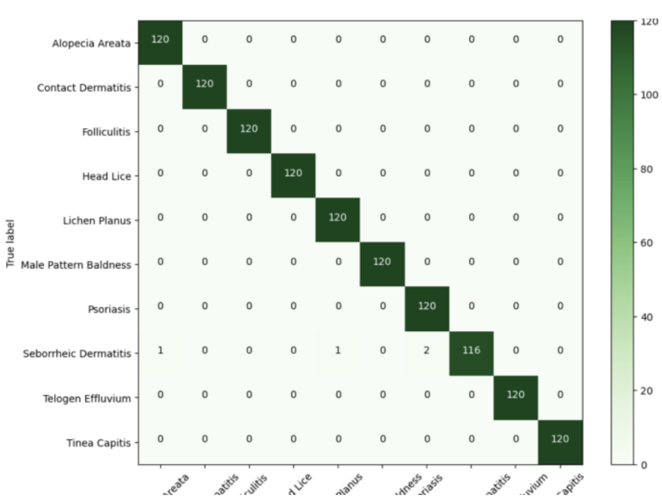
V. CONCLUSION AND FUTURE WORK:

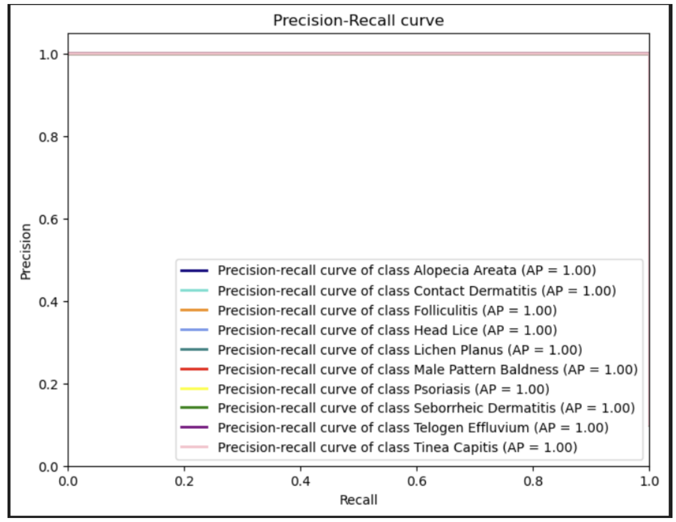
Throughout our research, we rigorously evaluated three state-of-the-art convolutional neural network architectures: MobileNetV2, InceptionV3, and DenseNet169, each pre-trained on the ImageNet dataset and fine-tuned for the classification of dermatological conditions. Our findings reaffirm the robustness of transfer learning in medical image analysis, with each model exhibiting commendable predictive capabilities. MobileNetV2's efficient architecture provided a delicate balance between speed and accuracy, making it a viable option for real-time applications. InceptionV3's inception modules facilitated the capture of complex features with its varied kernel sizes, leading to a profound understanding of dermatological patterns. DenseNet169 showcased exceptional performance with its densely connected layers, achieving an outstanding test accuracy of approximately 99.67%



38/38 ━━━━━━━━━━━━━━━━━━━━ 49s 1s/step

	precision	recall	f1-score	support
Alopecia Areata	0.99	1.00	1.00	120
Contact Dermatitis	1.00	1.00	1.00	120
Folliculitis	1.00	1.00	1.00	120
Head Lice	1.00	1.00	1.00	120
Lichen Planus	0.99	1.00	1.00	120
Male Pattern Baldness	1.00	1.00	1.00	120
Psoriasis	0.98	1.00	0.99	120
Seborrheic Dermatitis	1.00	0.97	0.98	120
Telogen Effluvium	1.00	1.00	1.00	120
Tinea Capitis	1.00	1.00	1.00	120
accuracy			1.00	1200
macro avg	1.00	1.00	1.00	1200
weighted avg	1.00	1.00	1.00	1200





Future Work:

Looking ahead, our team aims to broaden the scope of our research by expanding our datasets to include a more diverse range of skin conditions and demographic variables, thus enhancing the generalizability of our models. We are intrigued by the potential integration of multi-modal data sources, such as patient metadata and 3D image scans, to enrich the learning context for our models. Another avenue of interest is the exploration of interpretability frameworks for these deep learning models, providing clinicians with more transparent and explainable AI decisions. The ultimate goal is to establish a robust pipeline that not only excels in diagnostic accuracy but also seamlessly integrates into clinical workflows, thereby augmenting the capabilities of healthcare professionals and improving patient outcomes.

VI.: Contributions

Rithika performed the implementation and optimization of the MobileNetV2 architecture. Her tasks included configuring the model with pre-trained ImageNet weights, adjusting hyperparameters, training using Keras, evaluating performance metrics, and generating performance graphs for analysis.

Gousia took charge of tasks related to the InceptionV3 architecture, tailoring it for hair disease classification. This involved setting up the model with ImageNet pre-trained weights, fine-tuning layers, managing data preprocessing and augmentation, optimizing optimization parameters, conducting evaluations using F1-score, and creating visualizations like confusion matrices.

Harika led the development of the DenseNet169 model and managed data preprocessing. Her responsibilities encompassed configuring and training the model with pre-trained ImageNet weights, preprocessing tasks such as resizing and normalization, optimizing training with hyperparameters and dropout, validating performance with validation datasets, synthesizing results, and presenting detailed analytical reports and visual presentations.

Collaborative Tasks:

- **Model Integration and Final Review**: We collaborated to integrate the three models into a cohesive system, ensuring that each model's outputs and data handling processes are compatible.
- **Final Documentation and Presentation**: Each member contributed to the final project documentation, combining their sections into a comprehensive report that covers the methodology, results, discussion, and future work.
- **Peer Review**: Team members conducted peer reviews of each other's work to ensure accuracy, coherence, and technical integrity before the final submission.

VII. References

1) Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, 115–118. doi:10.1038/nature21056

2) Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA, 316(22), 2402-2410. doi: 10.1001/jama.2016.17216.

3) Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. [page numbers here], 2019).

4) Lin, R. L., Garibyan, L., Kimball, A. B., & Drake, L. A. (2016). Systemic causes of hair loss. Annals of Medicine, 48(6), 393-402. https://doi.org/10.1080/07853890.2016.1180426

5) Wolff, H., Fischer, T.W., & Blume-Peytavi, U. (2016). The Diagnosis and Treatment of Hair and Scalp Diseases. Dtsch Arztebl Int, 113(21), 377-386. https://doi.org/10.3238/arztebl.2016.0377. PMCID: PMC4908932, PMID: 27504707.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4908932/

6) Annalamai, S. Hair Diseases Dataset. Kaggle. https://www.kaggle.com/datasets/sundarannamalai/hair-diseases

7) Ahmed, A., Almohanna, H., Griggs, J., & Tosti, A. (2019). Genetic Hair Disorders: A Review. Dermatology and Therapy, 9, 421-448. https://doi.org/10.1007/s13555-019-0313-2

8) Roy, M., & Protity, A. (2022). Hair and Scalp Disease Detection using Machine Learning and Image Processing. https://www.researchgate.net/publication/366821677_Hair_and_Scalp_Disease_Detection_using_Machine_Learning_and_Image_Processing. doi:10.48550/arXiv.2301.00122