Herbal Hair Care

Ms. Anusha S J¹, Ms. Samyukta D Kumta²

¹Master of computer application PES University Bengaluru, India
²Department of Computer ApplicationsPES University Bengaluru, India

Abstract
Hair-loss and Scalp related issues are affected for many people due to factors such as stressing, medication, genetics, heredity, hormonal changes, differences between regular hair fall underline conditions often presents challenges, leading to delayed diagnosis and potentially more serious health issues. To address this issues, using neural network-based applications, which have revolutionized healthcare and health informatics. This study focuses on predicting three primary scalp disease are identified such as: Psoriasis, Alopecia areata and folliculitis and hair loss stages from stage 0 to stage 4 to detect this utilizing a convolutional neural network (CNN) and types using restnet50, vgg16, vgg19, efficientnet. Using this architecture to comparison between this types to predict accuracy model, help to more confidence particular disease effected and healthcare providers to offer personalized treatment plans and preventive measures tailored to each patient's unique circumstances. By addressing this hair-loss and scalp related issues by seeing this result we can go through dermatology by this we can reduce time and in this project any patient can go with testing and they can go with suggestion what they can do. By addressing this helps to reduce cost as this product is home remedies, this is natural remedies their will be won’t harm any chemical side effect. The model aids in preventing condition progression, ultimately reducing the strain of healthcare organization and this can lead ultimately better quality of life of users.

Keywords: Scalp related disease such as Alopecia areata (AA), Psoriasis, and Folliculitis and hair-loss, Dermatology, Convolutional Nural Networks(vgg16, vgg19, efficientnet, resnet50), MYSQL, ImageDataGenerator.

1. INTRODUCTIONS:
An average people loss(60-100) strands hair-loss in everyday sack is quite natural. The hair fall might happen when scalp density increases regarded as illness. Scalp disease significantly increases now a days by highly air pollution and dust particles leading to stages where every people can see scalp through the hair[1]. At the present time many people suffer from hair-loss and scalp related issues. In India 150 million people suffer from hair-loss, baldness, grey hair etc, 63% related to hair-loss and dandruff, 43% related to irregular sleeping pattern, 68% high stress, 45% weak metabolism etc they are other way we can see in order to get imbalance issues hair-loss and scalp disease[2]. Some of scalp related issues such Alopecia areata (AA), psoriasis, and folliculitis are a few scalp conditions that can cause non-scarring hair loss. The hair loss can affect a specific area of the scalp, the entire scalp, or the entire body[3]. AA it approximately 0.7% cases in India, psoriasis is baldness which absence hair in the head[4]. Psoriasis it approximately 0.44-2.8% cases in India, psoriasis which causes red itchy, scaly patches due to more which effect long-term with no-cure[5]. Folliculitis it approximately between 36.6% and 78.4% cases in India, Folliculitis it is bacteria infections like red pimple which appear in hair and...
In this paper, we can... otherwise require... for prevention such as scalp hair physiotherapy have been developed to treat various scalp disorders [4]. Physical therapy have been developed to address scalp problems... unbalanced diet, stress, and environmental toxins. Additionally, specific treatments such as scalp conditions such as alopecia areata, psoriasis, folliculitis, and healthy scalp[3]. These issues are caused by factors such as poor daily routine, unbalanced diet, stress, and environmental toxins. Additionally, specific treatments such as scalp physical therapy have been developed to address scalp problems. Recently, specific therapies designed for prevention such as scalp hair physiotherapy have been developed to treat various scalp disorders [4].
In Bangladesh, a machine learning-based study was conducted to diagnose hair loss and meningitis. By analyzing scalp images with different characteristics, the study proposed a classification system for conditions such as dry scalp, psoriasis, folliculitis, and healthy scalp regions.[5] Using Convolutional Neural Networks (CNN) to extract disease and hair loss using pretrained model resnet 50 and comparing vgg16, vgg19, efficientnetBO and transfer learning the accuracy rating were 92.64% get in the hair loss and checking rating in the scalp disease is 85.64% is get in hair loss stages. These accuracy data show how effective and dependable the suggested classification strategy is for classifying three distinct sets of scalp pictures and a four-stage hair loss pattern. Future research can use deep learning methods, including CNNs have undergone continuous research and refinement, leading to updated architectures and techniques that have further improvement in aims who have not grow hairs between 45 to more and some of genetics. Thus, utilizing a modern and well-designed CNN architecture, along with appropriate training methodologies, Convolutional Neural Networks (CNNs) is widely used in dl (deep learning) for tasks such as image recognition, object detection,[6] and image classification, among others. To do this, we created an image-based Hamilton-Norwood scale classification system for alopecia and three distinct scalp diseases, including psoriasis, folliculitis, and alopecia.[7] This dataset have been using unstructured by giving label making structured by easy to identifying the image reducing help lessen the effects of excessive fitting. It is evident from completed experiments that hair loss and scalp illness may be estimated from facial photographs utilizing deep learning architectures[8]. Potential study topics include fine-grained face centered image segmentations, data scaling, image generators, and head cropping as data enrichment techniques. Combining the test sets and selecting the best predicted class for each of them could be the subject of future research. [9] The proposed system utilizes a device equipped with dl (deep learning) capabilities to analyze scalp diseases, hair-loss patterns, and related conditions using real-time, high-resolution photos captured through a zoom-enabled camera particular part.[10] Additionally, a user-friendly web application interface facilitates interaction with the system, allowing individuals to upload their photos and receive accurate diagnostic results promptly and giving testing challenge to people trough this suggesting them in home remedies manner. Moreover, advancements in medication development targeting multiple factors contributing to hair loss offer promising avenues for improving treatment efficiency and achieving positive clinical outcomes.

3. RESEARCH METHODOLOGY:

Dataset: The most challenge aspect of utilizing visual pictures for sickness prevision and categorization be collection of data! We collected over 1300+ images for scalp diseases, categorized into three classes: psoriasis, folliculitis, and alopecia. Additionally, we gathered a dataset of over 700+ images for hair loss patterns, consisting of 4 stages: Stage 0, Stage1, Stage2, Stage3 ands Stage4.[11] Obtaining an adequate number of suitable images for specific illnesses can be difficult, especially considering that the images are often scattered across the internet. [12] In this study, the authors extracted images from various websites and other resources, including both native sources and real-time updated images.

3.1.1 Alopecia dataset: The dataset exclusively focuses on alopecia areata. It consists of over 539 images collected through web scraping and sourced from the Dermnet dataset, an online medical knowledge resource containing data on various dermatological disorders, including acne, eczema, and alopecia areata [13]. Figure 1 presents a selection of the images included in the alopecia areata dataset.
3.1.2 Psoriasis Disorder Dataset: The dataset exclusively focuses on psoriasis disorder. It comprises over 554 images collected through web scraping and sourced from the Dermnet dataset, an online medical knowledge resource containing data on various dermatological disorders, including acne, eczema, and psoriasis [14]. Figure 2 presents a selection of the images included in the psoriasis disorder dataset.

3.1.3 Folliculitis dataset: Folliculitis Dataset: The dataset exclusively focuses on folliculitis. It consists of 456 images collected through web scraping and sourced from the Dermnet dataset, an online medical knowledge resource containing data on various dermatological disorders, including acne, eczema, and folliculitis [15]. Figure 3 presents a selection of the images included in the folliculitis dataset.

3.1.4 Hair-loss dataset Stages: The dataset exclusively focuses on Hair-loss stages. It consists of 734 images collected through web scraping and sourced from the Dermnet dataset, each divided into 5 stages such as stage 0, stage 1, stage 2, stage 3, stage 4 an online medical knowledge resource containing data on various dermatological hair loss condition Figure 4 presents a selection of the images included in the hair-loss dataset.

3.2 Data Preprocessing:
3.2.2 Gaussian Blur: is a technique widely used in imaging to reduce image noise and detail. The image is convolved using a Gaussian kernel in order to create a smoothed version of it. The formula for Gaussian blur is given by:

\[ G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]

where \(x\) and \(y\) are the distances from the kernel's center, \(\sigma\) is the SD(standard deviations), and \(G(x,y)\) identified Gaussian kernel.

3.2.2 Median Blur: Median blur is a non-linear filtering technique used to remove noise from images. The value of each pixel is substituted with the neighborhood's median in the intensity value.
The formula for median blur is:

$$M(x, y) = \text{median}(I(x - k, y - k), I(x - k, y), ..., I(x + k, y + k))$$

Where $M(x, y)$ is the median filtered value at position $(x, y)$, $I$ is the input images, and $k$ is the kernel size.

### 3.2.2 Convert to HSV:

Converting an image from RGB to HSV color space separates the image into its Hue, Saturation, and Value components. This is a common preprocessing step in image analysis tasks. HSV separates the intensity information from the color information, making it easier to manipulate color values independently. The formula for converting RGB to HSV is:

$$\text{HSV} = \text{RGB} \times \text{RGB2HSV}$$

Where $\text{RGB}$ is the input RGB image and $\text{RGB2HSV}$ is the conversion matrix.

### 3.2.3 Data Augmentation:

Applying various transformation techniques to real photos to create altered duplicates of the original is known as image augmentation. By doing this, deep learning models can be trained on a larger variety of images than those found in the dataset. There is a chance of getting overfitting due to small dataset some changes of problem of corrupted therefore increase dataset augmentation techniques like resize rotation resolution. To implement picture augmentation in Keras, utilize the `ImageDataGenerator` class. The ability to generate real-time picture augmentation is a significant advantage of the Keras `ImageDataGenerator` class. This simply means that it can produce additional photos on-the-fly as the model is being trained, increasing the overall robustness and accuracy of the model.
4. DEVELOPING THE PREDICTION MODEL:

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 2</td>
<td>67</td>
</tr>
<tr>
<td>Stage 3</td>
<td>1</td>
</tr>
<tr>
<td>Stage 4</td>
<td>0</td>
</tr>
</tbody>
</table>

4.1 ResNet-50 is a CNN model used for images data recognition that belongs to the family of Residual Networks (ResNets). ResNets are renowned for their ability to train very deep neural networks effectively by addressing the vanishing gradient problem.

The architecture of the ResNet-50 consists of six components. Enter Data Prior to processing Cfg[0] Cfg is blocked[1]. Cfg is blocked[2]. Cfg is blocked[3]. Blocks fully Completely linked layer
Instead of utilizing the fully connected layers of the original ResNet50 architecture, I have replaced them with two Dense layers. These Dense layers are added on top of the output of the ResNet50 model to perform classification. Following steps given below:
Initialization: Load the pre-trained ResNet50 model without the fully connected layers (include_top=False).
Freezing Layers: Set the pre-trained model's layers as non-trainable to retain the learned features.
Adding Fully Connected Layers: Add custom dense layer on the top of the pre-trained model for fine-tuning.
Compile Model: For multi-class classification, it compiled the model using the Adam optimizers with calculating categorical cross-entropy loss.
Image Preprocessing: Preprocess the input image using the preprocess_input function specific to the ResNet50 architecture.
Prediction: Feed the pre-processed image into the model for prediction.
Formula: SoftMax Activation: The SoftMax activation function is applied to the output layer to convert raw scores into probabilities for each class. It computes the probability distribution over the classes.

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum e^{x_j}}$$

By implementing resnet50 In my project I’m getting accuracy 92.68% in scalp classification and 85.65% in hair loss stages

4.1.1 Table:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alopecia</td>
<td>99</td>
</tr>
</tbody>
</table>
Folliculitis 0
Psoriasis 0

4.1.2 table: Similar coding different dataset:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage0</td>
<td>21</td>
</tr>
<tr>
<td>stage1</td>
<td>9</td>
</tr>
</tbody>
</table>

4.1.3 Pseudo code:
Load pre-trained ResNet50 model without fully connected layers Add custom dense layers on top for fine-tuning
The model is constructed using the Adam optimizer and categorical cross-entropy loss.
Preprocess input image using ResNet50-specific preprocessing function
input pre-processed image into model for prediction.

4.2 vgg16 is a CNN model used for image recognitions it is a unique model which having 16 layers it relying on large number of hyper-parameters.

Vgg16 is a older model , It used for baseline for image classification task due to straightforward architecture is the best performance various simple datasets well suite for simplicity and interpretable.by implementing vgg16 I’m getting 90.31% in scalp classification and 88.26% in hair-loss detection.

4.2.1 table:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alopecia</td>
<td>87</td>
</tr>
<tr>
<td>Folliculitis</td>
<td>7</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>5</td>
</tr>
</tbody>
</table>

4.2.2 table:
Similar coding different dataset:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage0</td>
<td>99</td>
</tr>
<tr>
<td>stage1</td>
<td>0</td>
</tr>
<tr>
<td>stage2</td>
<td>0</td>
</tr>
<tr>
<td>stage3</td>
<td>0</td>
</tr>
<tr>
<td>stage4</td>
<td>0</td>
</tr>
</tbody>
</table>

4.2.3 Pseudo code:
Load pre-trained vgg16 model without fully connected layers Add custom dense layers on top for fine-tuning The model is constructed using the Adam optimizer and categorical cross-entropy loss.
Preprocess input image using vgg16-specific preprocessing function input pre-processed image into model for prediction.

4.3 vgg19 is a convolutional neural network model is a deeper architecture same as vgg16. It’s having 19 layer. Best performance then vgg16, particularly dealing with high dataset then vgg16. It used for benchmark in the field of computer vision suitable for task where interpretability and accuracy are prioritized over computational efficiency. Implementing vgg19 I’m getting 89.01% in scalp classification and 90.0% in hair-loss detection.

4.3.1 table:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alopecia</td>
<td>79</td>
</tr>
<tr>
<td>Folliculitis</td>
<td>18</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>2</td>
</tr>
</tbody>
</table>

4.3.2 table: Similar coding different dataset:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage0</td>
<td>0</td>
</tr>
<tr>
<td>stage1</td>
<td>36</td>
</tr>
<tr>
<td>stage2</td>
<td>62</td>
</tr>
<tr>
<td>stage3</td>
<td>0</td>
</tr>
<tr>
<td>stage4</td>
<td>0</td>
</tr>
</tbody>
</table>

EfficientNetB0 is a baseline version it is a family of convolutional neural network architecture that are designed to achieve state-of-the-art accuracy with fewer parameters.
and less computational cost compare to other models. It used for compound scaling methods that scales the depth, width and resolution of the network simultaneously achieve better efficiency, scales this dimensions in a coordinated manner to find optimal balance between model complexity and efficiency. EfficientNet being more resource-efficient compared to other architecture. By implementing VGG16 I’m getting 91.88% in scalp classification and 90.43% in hair-loss detection.

4.2.1 table:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alopecia</td>
<td>34</td>
</tr>
<tr>
<td>Folliculitis</td>
<td>59</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>5</td>
</tr>
</tbody>
</table>

4.2.2 table: Similar coding different dataset:

<table>
<thead>
<tr>
<th>Class</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage0</td>
<td>81</td>
</tr>
<tr>
<td>stage1</td>
<td>10</td>
</tr>
<tr>
<td>stage2</td>
<td>5</td>
</tr>
<tr>
<td>stage3</td>
<td>3</td>
</tr>
<tr>
<td>stage4</td>
<td>0</td>
</tr>
</tbody>
</table>

4.3.3 Pseudo code:

Load pre-trained EfficientNetB0 model without fully connected layers. Add custom dense layers on top for fine-tuning.

The model is constructed using the Adam optimizer and categorical cross-entropy loss.

Preprocess input image using EfficientNetB0-specific preprocessing function

input pre-processed image into model for prediction

Comparing 4 architecture old one VGG16 and extension VGG19 is best for used in baselines for comparison newer model. Compared to more recent on resent and efficientnet both are the best limited computational resources want high efficiency resource with good. This web application is built using Flask, a Python web framework. It incorporates functionalities for predicting scalp diseases and hair loss conditions, user authentication, and registration. When users access the home page, they are presented with a login form (‘login.html’) first user should register the credentials. After that users can sign in with their user ID and password. Before login the application it checks information in database to make sure that they are already registered. If the login is successful, users can access the dashboard. Means the New users can sign up by providing their user ID and password. The application saves this information in a
database so they can log in later in any times. When knowing hair-loss case and scalp disease, Users can upload pictures of hair-loss or their scalp disease. The application will check the images using machine learning to analyze the pictures and predict the scalp disease or hair-loss stage. It then shows the prediction and offers advice suggestion to prevent the people who are suffer from hair-loss and scalp condition. When the user input the image it saves in csv file. The application saves the prediction results and user information in files for future use. The application uses advance neural technology that have been trained to recognize different diseases and hair loss conditions. In these application I have compared 4 types of CNN like vgg16, vgg19, eficientnetBO, resnet50 comparing this and seeing the results and information the rsenet50 is the best comparing other type I have used resnet50 in my application it performs well. These application user friendly, user can input present real pic it checks and give results and seeing results it gives suggestion to people and product wise. The application is easy to use and helps users understand their hair-loss and scalp disease in the hair health. Additionally, In the application, user can see updated graphs of their hair loss stages, which can help them track their progress of accurate hair and make informed decisions about their health.

6. RESULTS AND DISCUSSIONS:

We separate the data into training and testing sets before the training a model. The model is taught by the training set, which gives it instances to work with. The testing sets is then used to the evaluate the model's performance and observe the outcomes it produces. We compares models based on precision, recall, F1-score, and accuracy. Precision tells us how often the model is correct when it predicts something positive, recall tells us how often the model finds all the positive things it should, and The F1-score combines both accuracy and recall.

To train the models, training and testing algorithms are developed from the data. The testing system evaluates the performance of the models in the training set. We compare the models based on four things: precision, recall, F1-score, and accuracy. Precision tells us how often the model is correct when it predicts something positive, recall tells us how often the model finds all the positive things it should, and the F1-score combines both accuracy and recall. We use equations to calculate these things, where true positive means the model predicted something correctly, false positive means the model predicted something incorrectly, and false negative means the model missed something it should have predicted.

\[
\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}
\]

\[
\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}
\]

\[
\text{F1-Score} = \frac{2(\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}
\]

To ensure a consistent evaluation of the model's performance, the recall value is calculated by dividing the number of true positive predictions by the sum of true positive and false negative predictions. Its recall and precision are combined to calculate the F1-Score. Although deep learning methods often provide better accuracy, they require more computing power and data. Because convolutional neural networks (CNNs) use filters instead of fully connected layers, they are particularly effective. Consequently, it is proposed to use CNN in future studies for the diagnosis of diseases such as acne, arthritis, psoriasis, and hair loss.
7. CONCLUSION AND LIMITATION AND FUTURE WORKS

7.1 Research Contribution: The project helps make healthcare technology better by using advanced neural network and image analysis to detect and manage hair loss and scalp disease problems by leveraging cutting-edge machine learning and image processing techniques to enhance the detection and management of hair loss and scalp diseases. By developing a user-friendly web application accessible to individuals worldwide, the project democratizes access to accurate diagnosis and personalized treatment recommendations, empowering users to take pre-emptive steps towards addressing their hair and scalp health concerns. The project's new way of doing things not only makes disease detection faster and more accurate but also gives people more options for treatment, especially natural remedies or cheaper treatments. The project is important because it helps more people access good healthcare and gives them more choices for taking care of their hair and scalp.

7.2 Limitations: The project has some limitations that might make it less helpful to the people and limit who can use it. The basic problem is that the application doesn’t help those who are more likely to lose all their hair and those who are age 45 years old and above might not work. This means that the application might not be able to help a wide range of people and might not be able to give complete care to all users. The accuracy of the application disease detection might also be affected in my project by things like quality of the image and resolution of the image and differences in scalp conditions. This could lead to mistakes, such as false positive or negatives, which might make users less confident in the application’s suggestions.

7.3 Future Work: To fix this problem with the project, we need to work on few things which are needed for the project. First, we need to make sure that our applications can help more people, including those who are more likely to lose their hair and those who are older and need to see what condition they are. We also need to keep working on the application's which are update technology to make it more accurate and reliable for all users, no matter their age or condition. Working with experts in hair and scalp health, such as dermatologists, ayurvedic, homeopathy and healthcare professionals, can help us give better advice and treatment options which we can get real time data and making our applications more easily and can save the time. By doing these things, we can make the application more helpful to all people and making innovative in the field of healthcare technology. We can also make sure that the application is more inclusive and this may get lot more people access good healthcare and gives them more choices for taking care of their hair and scalp.
of benefits a wider range of people.

REFERENCES


