Hair Product Recommendation System for Common Hair and Scalp Diseases

Joseph Yoo
Georgia State University
jyoo30@student.gsu.edu

Anirudha Narayanan
Georgia State University
anarayanan6@student.gsu.edu

Siya Katoch
Georgia State University
skatoch1@student.gsu.edu

Abstract

Advancements in machine learning and computer vision have paved the way for sophisticated algorithms and models in medical diagnosis and treatment. This project focuses on leveraging convolutional neural networks (CNNs) to predict three common hair and scalp conditions from images and recommend appropriate treatments. These conditions include alopecia, psoriasis, and seborrheic dermatitis. Such work can assist those affected in early detection and treatment planning, as such conditions often progress without concern or the ability to visit a professional for a diagnosis. The methodology includes data collection from diverse sources, preprocessing of the data, the testing of three popular pretrained convolutional image classification models, and a recommendation of treatment, given that a condition was predicted. The three models include ResNeXt, GoogLeNet, and AlexNet, and are used for our application, either through fine-tuning the models or using the models as a fixed-feature extractor. Augmentation using RandAugment was applied during training to expand the dataset and improve the generalization of the models. Finally, a survey of treatments for the hair and scalp conditions was conducted and used to design a decision tree that could be used to recommend treatments given the prediction of a condition. The limited availability of accessible and uniform data made our task of classifying hair and scalp conditions difficult and made our task of learning treatments according to hair and scalp conditions infeasible. Across the three models and our tests, we achieved a maximum validation accuracy of 99% and accuracies of 94.6%, 100%, 90.3%, and 97.8% for predicting alopecia, healthy, psoriasis, and seborrheic dermatitis, respectively, overall data.

1. Introduction

Hair diseases can adversely impact an individual's quality of life, specifically their comfort and self-esteem. Identifying hair products relevant to specific hair diseases is crucial for providing appropriate treatment. This project proposes to leverage convolutional neural networks (CNNs) and modern deep learning techniques to determine a person's hair or scalp condition, or lack thereof, and recommend hair products to treat their condition appropriately. This hair product recommendation will not be based on their condition but on the latent representation of the condition that our network produces. By doing this, we hope to address the issue of treatment based on the severity or variation of the condition, rather than treatment based on the label alone.

2. Related Work

Health evaluation is a long-time application of machine learning. Particularly, research into scalp damage recognition and hair disease identification using deep learning has begun relatively recently (2018, [5]). In a study by [1], physical and chemical analysis was conducted on 15,000 microscopic hair surface images to evaluate the degree of hair damage. A CNN, named SACN-Net, was then trained with 80% of these evaluations to predict the degree of hair damage based on the hair surface images. SACN-Net utilizes residual connections, attention modules, and a global average pooling and performed better than several other CNN models, including ResNet50, with a test accuracy of 98.38% on 20% of the data.

Another study, by [2], used 70% of a 150 scalp image dataset, collected from various sources, classified with alopecia, psoriasis, and folliculitis to train a CNN consisting solely of max pooling and convolution layers, with one fully connected (FC) layer before the output. Their model reached a "validation" (testing, in their case) accuracy of 91.1% on 30% of the data.

A holistic, robust system, called ScalpEye, for efficient inspection and diagnosis of four common scalp hair symptoms, was developed by [3], which consists of a portable imaging microscope, mobile app, training server, and cloud management platform. They tested several popular models before settling on the R-CNN Inception ResNet_v2_Atrocious model for image recognition. ScalpEye was experimentally shown to diagnose dandruff, folliculitis, hair loss, and oily hair with an average precision of 97.41% to 99.09%.

Overall, an assortment of work has been conducted on the evaluation of scalp health and identification of scalp hair diseases, with many having impressive results. However, no studies on hair product recommendation systems were found. In this work, we will leverage and build upon the work of others and utilize these techniques to construct an effective hair treatment recommendation system.

3. Methodology

3.1. Data Collection

A drawback of deep learning is the large amount of data required for an accurate, generalized model, especially when data is seldom available, labeled, and free.

For our hair and scalp condition classification problem, we collected images labeled with hair and scalp conditions and healthy hair. Images labeled with hair and scalp conditions were collected from Kaggle, a data request, and a Github repository, originating from the DermNet and DermNetNZ datasets. Images labeled with healthy hair were collected from the Figaro 1K dataset. The number of images for each category is listed in Table 1. A sample image from each category is shown in Fig. 1

Label	Quantity
Healthy	1050
Alopecia	147
Psoriasis	134
Seborrheic Dermatitis	90

Table 1: Images per Label

Unfortunately, we could not collect data for our hair treatment recommendation problem. As stated in our introduction, we hoped to administer a treatment based on a latent representation of the condition, rather than the label of the condition alone. To achieve this, we would require a dataset that includes images of hair and scalp conditions, labels of the condition, and prescribed treatments. Such a dataset was not readily available on the internet.

3.2. Data Processing

Data will be preprocessed according to conventional techniques for transfer learning. This process is shown in Fig. 2. The standard ImageNet transforms include a resizing to 256x256 pixels using bilinear interpolation, a central crop to 224x224 pixels, and a normalization to the mean and standard deviation of the dataset. Augmentation will be applied to diversify our limited training data during training using RandAugment. This technique chooses and applies some number of randomly chosen transformations at some severity. The number of transformations and their severity are two hyper-parameters of a hyper-parameter search problem based on validation accuracy [9].
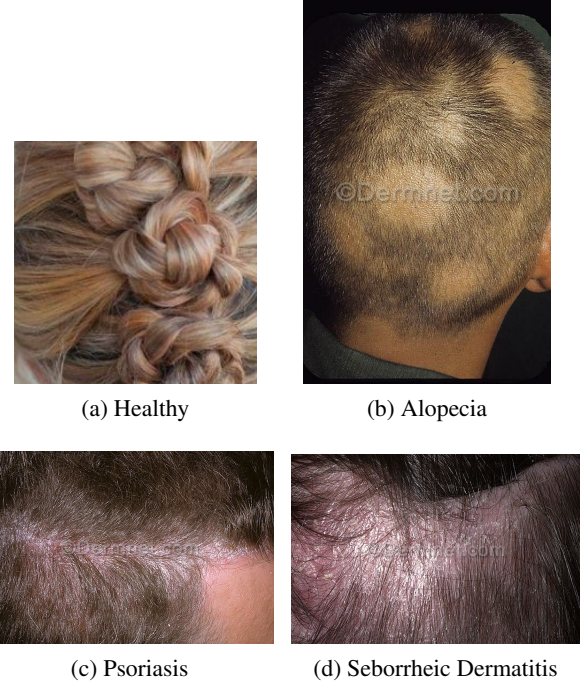

(a) Healthy   (b) Alopecia

(c) Psoriasis   (d) Seborrheic Dermatitis

Figure 1: Sample of Each Category

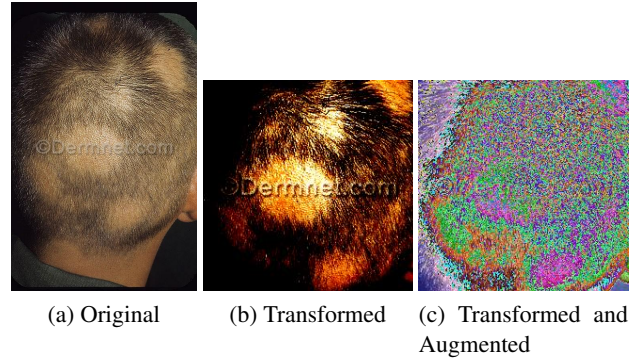(a) Original   (b) Transformed   (c) Transformed and Augmented

Figure 2: Image Preprocessing

3.3. Model

We will apply transfer learning to three popular pre-trained models to classify hair and scalp conditions. These models include AlexNet [12], GoogLeNet [10], and ResNeXt [7], which all utilize convolutions. AlexNet is a classic convolution network that pushed the capabilities of CNNs at the time. GoogLeNet presented the idea of "Inception Modules", which could be stacked to build a robust network [10]. Finally, ResNeXt builds upon another model, ResNet, by exploring cardinality in addition to depth and width. Cardinality refers to the number of paths taken through a residual block. By increasing the cardinality, ResNeXt models could outperform ResNet models of sim-

ilar complexity [7]. Each model will be trained, and the best-performing weights of each model will be saved and compared. The best-performing model will be used for the condition classifier of the hair and scalp treatment recommendation system.

For transfer learning, we can either fine-tune the model or use the model as a fixed-feature extractor. Considering that the new dataset is small and different from the original dataset (ImageNet), we should use the model as a fixed-feature extractor to avoid over-fitting [11]. However, we will try both techniques and compare their performances.

One of the factors that the generalization of a deep learning model depends on is the balancedness of the dataset. A great imbalance risks the model developing a bias for the majority label(s). If so, a model may exhibit a high accuracy, but a confusion matrix will reveal that the model performs well only on the biased labels. Given that most of our data is labeled as "Healthy", our model will likely develop a bias for that label. A variety of techniques exist to circumvent this issue. We will compare three techniques: a control, weighted random sampling, and a modified loss function. In the control, batches will be sampled with the assumption that the data is balanced. Weighted random sampling is randomly sampling from the data with a stronger emphasis on the under-represented data and a weaker emphasis on the over-represented data. These weights can also be thought of as the inverse of the frequency of each label. The modified loss function is a multi-class implementation of a technique called focal loss [13], whose form is shown below:

$$\text{FL}(p_t) = -\alpha_t(1 - p_t)^\gamma \log p_t \tag{1}$$

Simply put, it can be thought to "encourage" the model to prefer the predictions of false positives over false negatives (i.e. predicting a condition when there is none is preferable to predicting no condition when there is one).

3.4. Treatment Recommendation

Despite our inability to collect data relevant to this part of our problem, we collected treatment options for the scalp and hair conditions of this work into a decision tree. This decision tree is shown below in Fig. 3
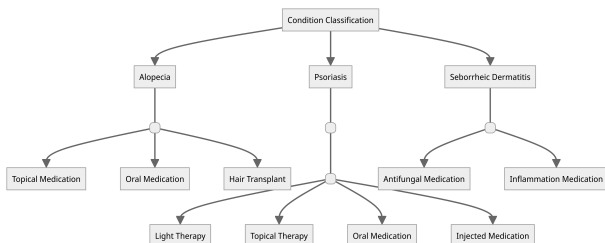

Figure 3: Decision Tree of Treatment Options

4. Experiments and Results

Across the three models and transfer learning techniques, we found that fine-tuning the pretrained GoogLeNet to our data yielded the best validation accuracy. Table 4 shows the maximum validation accuracies of all models.

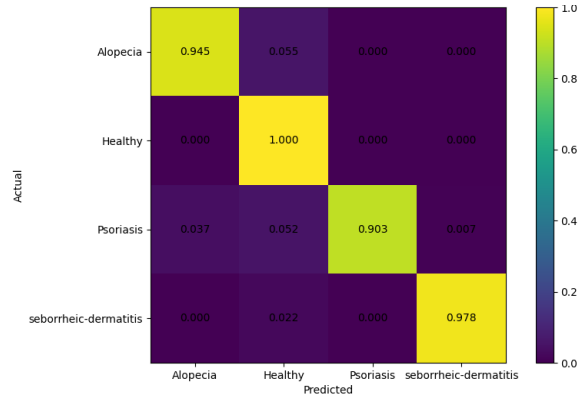| ResNext-50	96%
GoogLeNet	99%
AlexNet	90%

Table 2: Validation Accuracies of Each Model

Figure 4: Confusion Matrix of Fine-tuned GoogLeNet

However, the confusion matrix of the fine-tuned GoogLeNet matches what we hypothesized, that the imbalance of data would cause the model to bias the over-represented data. This disparity is seen in Figure 4, where healthy samples are predicted with 100% accuracy while psoriasis samples are predicted with only 90.3% accuracy. To mitigate this, we tried applying weight random sampling and applying focal loss to our training. We were willing to suffer an overall accuracy to gain a reduction in the variance of the accuracies across the labels. Surprisingly, neither method was effective. Weighted random sampling led to over-fitting across multiple trials, resulting in a training accuracy of 100% and a validation accuracy of 75%. In addition to significantly longer training times, applying focal loss gave an overall accuracy worse than the control, with a maximum validation accuracy of 93% and a more severely biased model. As shown in Figure 5, our model trained with focal loss often misclassified psoriasis, correctly classifying only 59% of samples.

5. Conclusions

Hair and scalp conditions require quick diagnosis and treatment for effective mitigation. Thus, a lack of awareness, lengthiness of diagnosis, and cost of diagnosis and treatment all stand in the way of effective mitigation. With the context of prior work relevant to our problem in mind,
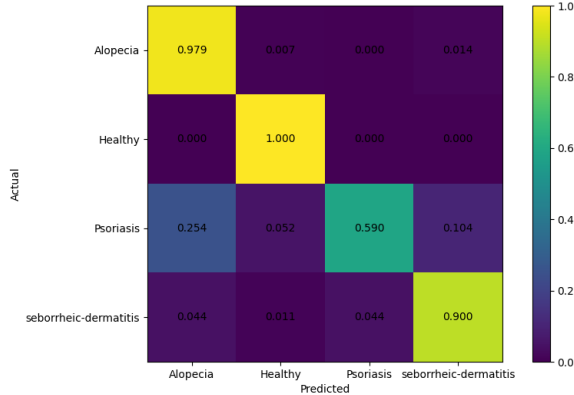
Figure 5: Confusion Matrix of Fine-tuned GoogLeNet with Focal Loss Applied

we applied transfer learning to three powerful pre-trained models to classify three common hair and scalp conditions, and the lack thereof, using 371 images of hair and scalp conditions, and 1050 images of healthy hair. After training, we achieved a high maximum overall validation accuracy of 99% using GoogLeNet. However, our model presented biases that our techniques were unable to mitigate, predicting healthy hair with extremely high accuracy (100%) and predicting psoriasis with a significantly lower accuracy (90.3%). Furthermore, an absence of data correlating hair conditions or images of hair conditions with treatments rendered our goal of learning treatment predictions infeasible.

References

[1] Qiaoyue Man, Lintong Zhang, and Youngim Cho. Efficient Hair Damage Detection Using SEM Images Based on Convolutional Neural Network. *Applied Sciences*, 11(16):7333, Jan. 2021. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute.

[2] Mrinmoy Roy and Anica Tasnim Protity. Hair and Scalp Disease Detection using Machine Learning and Image Processing, May 2023. arXiv:2301.00122 [cs].

[3] Wan-Jung Chang, Liang-Bi Chen, Ming-Che Chen, Yi-Chan Chiu, and Jian-Yu Lin. ScalpEye: A Deep Learning-Based Scalp Hair Inspection and Diagnosis System for Scalp Health. *IEEE Access*, 8:134826–134837, 2020. Conference Name: IEEE Access.

[4] Hrishabh Tiwari, Jatin Moolchandani, and Shamla Mantri. Diagnosis of Scalp Disorders using Machine Learning and Deep Learning Approach – A Review, Aug. 2023. arXiv:2308.07052 [cs].

[5] Ishita Kapoor and Anju Mishra. Automated Classification Method for Early Diagnosis of Alopecia Using Machine Learning. *Procedia Computer Science*, 132:437–443, Jan. 2018.

[6] Halim Benhabiles, Karim Hammoudi, Ziheng Yang, Feryal Windal, Mahmoud Melkemi, Fadi Dornaika, and Ignacio 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)*, pages 1–6, Nov. 2019. ISSN: 2154-512X.

[7] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated Residual Transformations for Deep Neural Networks, Apr. 2017. arXiv:1611.05431 [cs].

[8] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[9] Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, and Quoc V. Le. Randaugment: Practical automated data augmentation with a reduced search space, 2019.

[10] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions, 2014.

[11] Fei-Fei Li, Justin Johnson, and Serena Yeung. Cs231n: Convolutional neural networks for visual recognition, 2024.

[12] Alex Krizhevsky. One weird trick for parallelizing convolutional neural networks, 2014.

[13] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection, 2018.