JOURNAL OF
EMERGING INVESTIGATORS

Article

Quantitative analysis and development of alopecia areata classification frameworks

Ayushmaan Dubey[1], Arturo Morales[1]
[1] Meridian World School, Round Rock, Texas

SUMMARY

Alopecia areata is an autoimmune disorder resulting in rapid and unpredictable hair loss on the scalp or body as the immune system mistakenly attacks human hair follicles. In the United States alone, about 6.7 million people experience a form of Alopecia. Early identification of the condition has shown notable potential in improving treatment outcomes and reducing complications. To diagnose Alopecia, researchers have proposed the use of deep learning (DL) techniques to classify images of hair as healthy or alopecia-affected, which has shown high potential. However, the research implementing relevant DL algorithms in the field of hair loss detection and estimation is limited. This paper presents a comparative analysis of our two newly optimized Convolutional neural networks (CNN) with other existing models. For training, we considered datasets comprised of images of healthy hair and alopecia-affected hair. Due to data unavailability, we gathered images of alopecia-affected hair from two distinct datasets: one from Figaro1k and one independently created dataset. After training the algorithms, we performed a contrastive assessment to determine the most effective one based on relevant criteria. We hypothesized that the initial performance of the base neural network would be closely connected to the subsequent accuracy of the algorithm when training it for a new task. As expected, the modified Inception-Resnet-v2 model achieved the greatest performance, with a validation accuracy and loss of 97.94% and 10.4%, respectively. The experimental results indicated that the proposed algorithm serves as an effective framework for Alopecia Areata classification.

INTRODUCTION

Alopecia Areata is an autoimmune disorder that produces non-scarring hair loss, typically in small visible patches, and ultimately affects the entire region where the patches initially occurred (1). As of now, scientists do not entirely understand the causes of this unintentional response. However, researchers believe that genetic and environmental factors are the primary causes for the condition (2). This condition is especially prevalent in individuals with a family history of Alopecia Areata. According to the National Alopecia Areata Foundation, about 160 million people worldwide have experienced alopecia areata, with 6.7 million people in the United States alone (3). The initial cause of the disease occurs when the body's autoimmune system targets the individual's hair follicles and prevents further hair growth (1). To ensure early and effective treatment for the patient, it is important to promptly and accurately identify Alopecia Areata. This would significantly reduce the potential for further hair loss.

Currently, in clinics, some common diagnostic methods for hair loss estimation include the hair pull test, the pluck test, a scalp biopsy, daily hair counts, or most commonly, trichoscopy (4). Trichoscopy involves examination of the scalp and hair using a handheld or a videodermoscopy device. It also appears to be the most effective method for diagnosis. Trichoscopy records the progression of hair loss in a given time period and evaluates hair condition and hair follicles based on certain metrics, including the actual quantity and diameter of total hairs. However, this method is not the most accurate diagnosis as it heavily relies on visual examination from a doctor, making it prone to human error. Additionally, many individuals find trichoscopy diagnosis extremely unaffordable as it costs about $14,000 USD (4).

Recently, computer-processed analyses using deep learning (DL) methods have aided in the prediction and classification of diseases. DL models are essentially multi-layered neural networks that mimic the learning processes of the human brain (4). When compared to traditional Machine Learning (ML) algorithms, DL algorithms exhibit excellent efficiency in predicting and classifying various diseases and disorders, as these frameworks are intended to extract nuanced features and meaningful relationships from images. This is possible through several hidden layers that act as processing methods located between the input layers, where the image data is received, and the output layers, where the image is classified. Object detection and image segmentation are two common techniques for DL classification. Object detection utilizes bounding boxes to specify the location and extent of identified objects. Image segmentation produces pixel-wise masks to precisely outline different regions of interest (ROI) within an image. Image segmentation is frequently applied alongside tasks like object detection, as it extracts fixed-size feature maps of an input image. This DL approach is most effective for algorithms that classify Alopecia Areata and estimate hair loss severity (4).

However, DL methods currently face some limitations, and there is a need to enhance their accuracy for more reliable diagnoses. Recent approaches exhibit irregular false positive rates, inaccuracies in detecting and classifying hairs, and generally provide erroneous hair loss approximations (4). Additionally, the lack of availability of relevant datasets and the high degree of variability among different images for the task pose challenges for developing this neural network. Finally, every image within a dataset requires different preprocessing

https://doi.org/10.59720/24-004

methods before algorithm training, especially if images differ in size and resolution (5).

In the field of dermatology, some publicly available DL algorithms have demonstrated relatively moderate-high accuracy for Alopecia Areata detection (**Table 1**). Some algorithms, for instance, ScalpEye, a multi-layered network for scalp inspection and diagnosis system implemented a Faster R-CNN— a relatively standard neural network that identifies regions of interest which are then passed to a second network that classifies the image— with the advanced, multilayered Inception ResNet_v2_Atrous model for image classification with high accuracy and an average precision (AP) ranging from 97.41% to 99.09% (6). A paper also introduced an unsupervised hair segmentation and counting system, which showed a precision rate of 95.3%; however, it has a manual parameter selection, making it difficult to operate. (7). Some researchers have proposed other unique classification frameworks, such as a DL algorithm that classifies patterned baldness from facial images. (8). Furthermore, an especially augmented combination of DenseNet, XceptionNet, and ResNet achieved an accuracy of 95.84% (9). There was also an EfficientDet algorithm with an accuracy of 81.74% (10).

Overall, the primary concern for these methods is their return frequency of higher false positive and false negative rates. The presence of false readouts either leads to unnecessary treatment, or an undetected disease. Some methods even present challenges regarding lots of image pre-processing and are not yet able to accurately classify a particular stage of Alopecia. Collectively, these recent approaches have exhibited favorable outcomes; however, the dermatology industry still needs a more robust, applicable algorithm that not only effectively classifies Alopecia Areata but also minimizes false positive and false negative rates during diagnosis.

To address these limitations, we developed two effective DL models using Transfer Learning (TL): Xception and Inception-resnet-v2. TL involves using a pre-trained model developed for a different task and modifying it to perform a new task. This technique has proven to be very useful as it utilizes pre-existing neural network knowledge for performance and reduces training time.

Consequently, the paper had two primary purposes: proposing two new, high-accuracy DL algorithms for Alopecia Areata detection, and synthesizing the evaluation on those trained models to determine the most effective framework for this task. In this paper, we demonstrate a realistic and robust DL algorithm that accurately classifies various stages of Alopecia Areata. The most optimal algorithm demonstrated a final accuracy of 97.94% and a final validation loss of 10.4%. The results from this paper display the strong potential of DL algorithms in dermatological settings.

Initially, we hypothesized that the neural network with the most effective base neural network, prior to TL modifications, would produce the best results because the pre-existing information in the original model would serve as a useful foundation for the modified neural network. In order to test this hypothesis, we comparatively analyzed the two algorithms based on a quantitative standpoint. We then compared the algorithms to some existing algorithms for Alopecia Areata classification.

During training, we considered key characteristics such as the presence of hair loss, the thickness of hair, number of bald patches, and overall hair loss severity. For clarity, initially, the selected models differed in terms of their accuracy and speed: Inception-resnet-v2 proved superior in accuracy while Xception in speed. Moreover, we chose only two models for this investigation, as these had the 2nd and 3rd highest accuracies among all pre-trained networks available in MATLAB, the numerical computing software used for this study. We used accuracy as the sole criterion to assess the algorithms.

DL Models	Final Validation Accuracy
Modified Inception-ResNet-v2 (This study)	97.94%
Modified Xception (This study)	96.71%
Mask R-CNN (4)	79.29%
DenseNet + XceptionNet + ResNet (9)	95.84%
Multi-Disease Detector (5)	91.1%
Attention-based Balanced Multi-Task Deep (AB-MTDeep) learning system (16)	95.11%
Attention-based Balanced Multi-Tasking Ensembling Deep (AB-MTEDeep) System (18)	96.94%
VGG-SVM (17)	98.31%

Table 1: Framework Comparison with other DL algorithms. The classification accuracies of our modified Neural Networks were compared to those of other relevant Alopecia Detection deep learning Frameworks.

Ultimately, this investigation not only provides a robust DL algorithm for accurately classifying the Alopecia Areata, but also presents an analytical assessment about the top-performing DL algorithms for this task. This paper demonstrates a comprehensive understanding of the critical application and relevancy of DL in the dermatological industry.

RESULTS

To construct high-accuracy DL algorithms for Alopecia Areata classification, we trained and utilized two base classification frameworks in this work: Inception-ResNet-v2 and Xception. We extracted ground truth data of Alopecia-affected hair and unaffected hair using the following mentioned datasets. For training, we collected a total of 1050 healthy hair images from the Figaro1k dataset (11). This is a publicly available dataset that contains a variety of hair images such as straight, wavy, and curly. It also contains images of healthy hairs from different ethnicities, making the algorithm more realistic in any clinical setting. For the unhealthy dataset, we extracted a total of 171 relevant Alopecia Areata images from two publicly available datasets: Kaggle, Dermnet, (12) and Scalp Disease detection (5), an independently created dataset. Then, we modified and trained Inception-ResNet-v2 (**Figure 1**) and Xception (**Figure 2**) using TL on MATLAB.

Inception-ResNet-v2

Researchers integrated two very successful algorithms, CNN's ResNet and Inception, to create the original Inception-ResNet-v2 (13). For network optimization on a substantial number of filters (exceeding 1,000), this network utilizes a method of efficiently diminishing the residual. This strategy effectively mitigates instability issues common in residual variants, which otherwise hinder network training when the filter count is greater than 1,000 (14). Ultimately, this

JOURNAL OF
EMERGING INVESTIGATORS

https://doi.org/10.59720/24-004

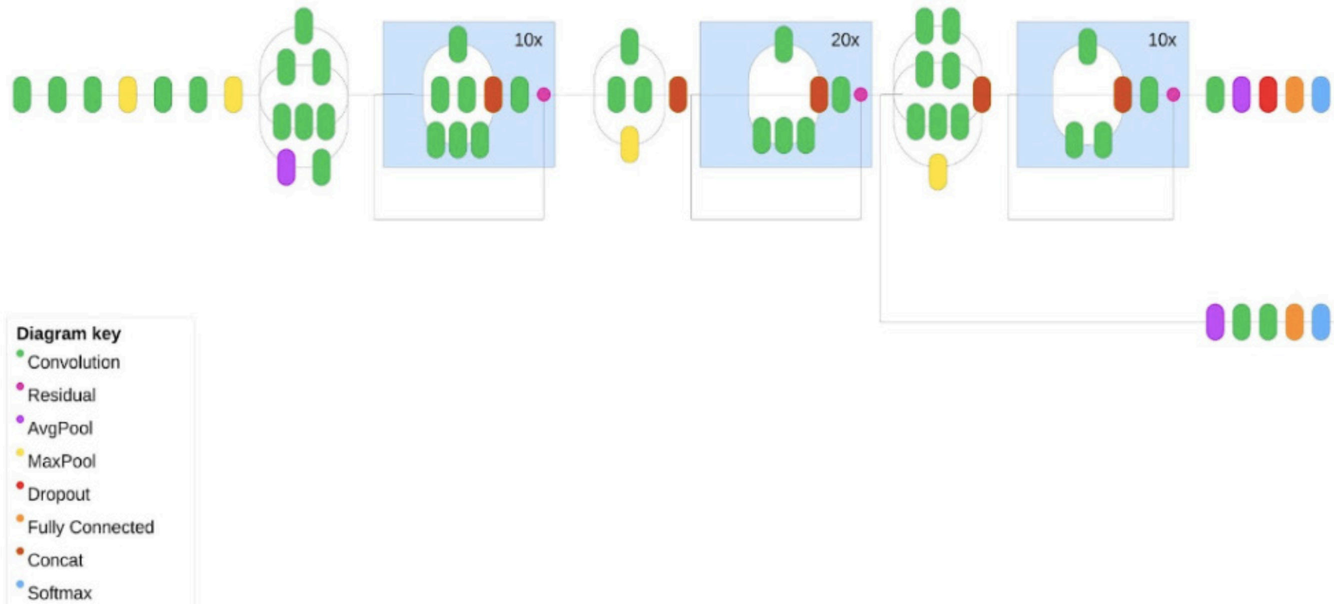

Figure 1: Figure 1. Schematic Diagram: Compressed View of Base Neural Network Inception-ResNet-v2. The authors used LucidChart software to model the Inception-ResNet-v2 architecture. The various blocks represent different components of the architecture: green represents convolution layer, pink for residual connections, purple for the average pooling layer, yellow for the max pooling layer, red for the dropout layer, orange for fully connected layer, brown for the concatenation layer, and blue for the SoftMax layer. The light blue boxes represent the sizes of the convolutional filters used in the depthwise separable convolution blocks.

characteristic aids in the stabilization of network training (**Figure 1**).

Xception

This original CNN architecture is structured around a linear stack comprising 36 depth-wise separable convolution layers (15). This network contains two important convolutional layers: a depth-wise convolutional layer that executes spatial convolutions independently in each channel of input data, and a pointwise convolutional layer, where a 1 × 1 convolutional layer transforms the output channels into a new channel space using depthwise convolution (15) (**Figure 2**).

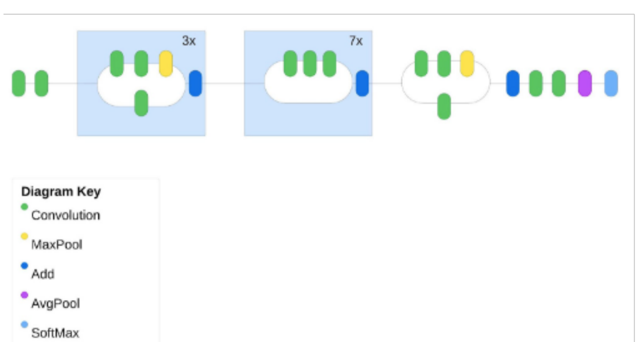

Figure 2: Schematic Diagram: Compressed View of Base Neural Network Xception. The authors used LucidChart software to model the Inception-ResNet-v2 architecture. The various blocks represent different components of the architecture: green represents convolution layer, yellow for the max pooling layer, dark blue for the add (addition) layer, purple for the average pooling layer, and blue for the SoftMax layer. The light blue boxes represent the sizes of the convolutional filters used in the depthwise separable convolution blocks.

We selected these algorithms because they achieved the highest accuracies among all the pre-trained networks provided by MATLAB. However, as mentioned previously, the device used for algorithm training has limited capabilities and simply cannot handle a network as complex as NASNET-Large, the most accurate base network provided by MATLAB. Furthermore, we considered the actual speed of the neural network irrelevant to the study, as the industry needs a more accurate algorithm. Thus, we did not consider it as a potential criterion for algorithm selection.

Performance and Evaluation

After training, we first found the final validation accuracy and loss for both neural networks. The validation accuracy refers to the accuracy percentage of the correctly identified instances, while validation loss quantifies the discrepancy between the predicted and observed values. These metrics are calculated through a separately assigned validation dataset.

The computer computes the accuracy for both algorithms by dividing the number of correctly classified images by the total number of test images and multiplying the result by 100. Formally, this is expressed by the following equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100$$

In this equation, TP represents true positives; TN is true negatives; FP is false positives; FN represents false negatives. Furthermore, to calculate the loss, we use the multi-label classification formula since the dataset is binary:

$$Loss = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{K}(T_{ni}\ln(Y_{ni}) + (1 - T_{ni})\ln(1 - Y_{ni})$$

JOURNAL OF
EMERGING INVESTIGATORS
https://doi.org/10.59720/24-004

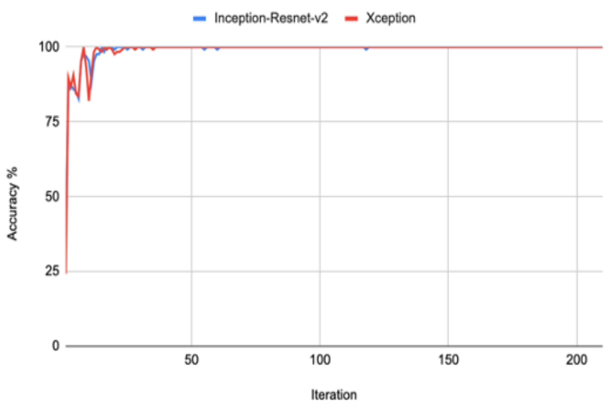

Figure 3: Comparative assessment of Classification Accuracy (%) for modified Inception-Resnet-v2 and Xception networks. A bivariate line graph was used to analyze the Classification Accuracy (%) for modified Inception-Resnet-v2 and Xception networks.

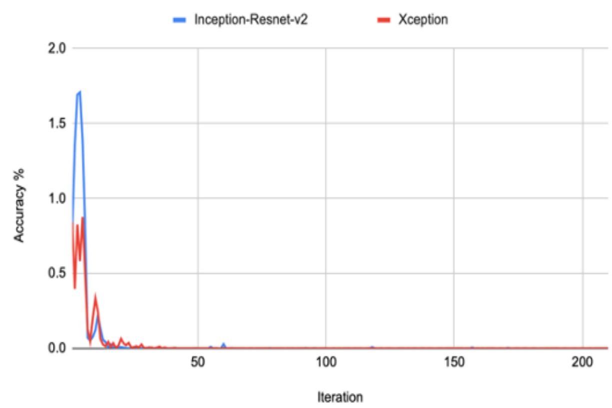

Figure 4: Comparative assessment of Classification Loss curves for modified Inception-Resnet-v2 and Xception networks. A bivariate line graph was used to analyze the Classification Loss (%) for modified Inception-Resnet-v2 and Xception networks.

Here, N denotes the number of observations, and K represents the number of classes. T_{ni} is the target value for observation n and class i. Yni refers to the predicted probability of observation n relating to class i.

The modified algorithms resulted in a final validation accuracy of 97.94% for Inception-ResNet-v2 and 96.71% for Xception. The networks achieved corresponding final validation losses of 10.4% and 11.3%, respectively. For context, in this investigation, we split the data 80/20, using 80% for the training dataset and 20% for the validation dataset. Unfortunately, MATLAB does not report the validation accuracy and loss for each iteration. Therefore, for interpreting the models' performance for every iteration, we closely analyzed training accuracies and losses. This is a reasonable approach because the only difference between the training and validation data is that while training data measures the accuracy and loss the model was trained on, and validation data measures both metrics on a separate, smaller, unseen dataset.

Ultimately, the data from this study suggests that both DL algorithms achieved a high detection accuracy. We interpreted that the Inception-ResNet-v2 algorithm achieved a higher training accuracy rate on average per iteration. However, both algorithms demonstrated a final training accuracy of 100% (**Figure 3**).

For both object detection algorithms, the loss curves reached a stable convergence in accordance with the hyperparameter configurations; this indicates that the models have effectively learned the information from the training data. Moreover, from a quantitative perspective, the graph indicates that Inception-ResNet-v2 exhibited a higher loss rate compared to Xception in the initial 20 iterations. Inception-ResNet-v2 achieved a final training loss of around 0.000318%, while Xception reached approximately 0.011637% (**Figure 4**).

Comparative Accuracy Performance of Previous DL Algorithms for Alopecia Areata Detection

Following evaluation, we compared the modified neural networks to other relevant Alopecia Detection Frameworks. We focused on assessing accuracy, as many publications did not provide information on the loss rates of the other networks. We decided not to analyze other metrics like mean average precision (mAP), recall, and the f1 score because these are not necessary in this situation. The dermatological industry needs a more accurate algorithm, and it is important to show how the learning performance of a model as this would reveal its overall learning efficiency for binary problems. Consequently, the only relevant metrics would be the loss and accuracy rates, as these criteria measure how well the algorithm can learn data and return correct information, respectively. Only publications that precisely listed the validation accuracy were considered in the analysis (**Table 1**).

The VGG-SVM network achieved the highest accuracy of 98.31% (17); however, it is important to note that this neural network had a significant limitation, as it did not account for racial differences. Therefore, it is non-applicable in many clinical settings. The modified Inception-RESNet-v2 algorithm from this study achieved the second highest accuracy. Next, an Attention-based Balanced Multi-Tasking, the Ensembling Deep (AB-MTEDeep) system achieved an accuracy of 96.94% (18). Through cross residual learning, researchers composed this algorithm of the Faster Residual Convolutional Neural Network (FRCNN) and Long Short-Term Memory (LSTM) network to classify scalp images. Thus, this algorithm achieved high accuracy by generating high-quality images and extracting confidence maps along with the bust depth maps from real hair and scalp images. Confidence maps represent the certainty level of a model's predictions, while bust depth maps represent the spatial depth information of that correspond to different points in an image. Our modified Xception network followed this network and achieved an accuracy of 96.71%.

DISCUSSION

The results from this study reveal that despite both the Inception-Resnet-v2 and Xception neural networks displaying similar accuracies and loss rates, Inception-Resnet-v2 achieved a higher accuracy and a lower loss rate. Specifically, Inception-Resnet-v2 exhibited a slightly higher accuracy of 97.94% compared to Xception's 96.71%,

JOURNAL OF
EMERGING INVESTIGATORS
https://doi.org/10.59720/24-004

along with a marginally lower loss rate of 10.4% compared to Xception's 11.3%. Therefore, we concluded that for the selected datasets, Inception-Resnet-v2 is a more effective algorithm for binary classification.

One of the primary reasons that Inception-Resnet-v2 proved to be the most effective algorithm was due to its depth and capacity. Greater depth and complexity in neural networks typically contribute to higher accuracies in image segmentation tasks, as the generated feature maps include more information. In the case of Inception-Resnet-v2, each layer captures specific facets of input data, and their sequential stacking significantly improves the model's discernment of distinct feature relationships. More specifically, the Inception-ResNet-v2 network integrates residual connections derived from the ResNet architecture which is often recognized for its efficacious implementation of these residual connections (15). These residual connections can be algebraically expressed by:

$$y = f(x) + x$$

where y is the final output of the module, f(x) represents the processed output and x is the initial input. The residual connections significantly reduce vanishing gradients and ensure a consistent flow of gradients during the backpropagation process. This characteristic allows the network to determine more abstract and hierarchical features from input images.

Another potential reason for the variation in accuracy and loss was the different feature extraction techniques applied by the algorithms. In Inception-Resnet-v2, convolutional layers in the first stem capture basic features of an image, while the convolutional processes inside the Inception blocks conduct multiscale feature learning. Moreover, the training difficulties in deep networks are significantly lowered by the use of residual connections (15). On the other hand, Xception uses depthwise separable convolutions for feature extraction. The Xception CNN architecture contains a series of these depthwise separable convolution blocks for capturing regions of interest within an image. This characteristic allows the network to efficiently extract complex image features. Additionally, this CNN also utilizes skip connections, which aids in gradient flow. (19). Although this technique allows for greater efficiency, high accuracy is not always guaranteed. After feature extraction, both functions enter the same Logistic Regression layer (SoftMax) which can be modeled by the equation:

$$\sigma(z)_i = \frac{e^{zj}}{\sum_{i=1}^{K} e^{zj}}$$

where is the output SoftMax value, z is the input vector, ezi represents the exponential function applied to the raw score for class i, K is the number of classes, and ezj is the exponential function for the output vector. The denominator is the normalization factor that takes the sum of exponentials of all raw scores.

Therefore, the higher accuracy attained by Inception-ResNet-V2 can be reasonably attributed to the combination of inception blocks and residual connections, as these enable the accurate capture of multi-scale features. The final SoftMax layer remains consistent between both networks, thus, it does not alter the outcome after the feature extraction technique.

It is also important to note that the dataset was imbalanced. In order to train the networks, we used 1050 images from a healthy hair dataset and 171 images from an Alopecia Areata dataset. Hence, we assigned approximately 14% of the total data as alopecia-affected images. Typically, this imbalance would suggest the presence of too many false positive or false negative results, and a 50/50 split is generally recommended. This would often mean that the accuracy rate will decrease, and the loss rate will increase. However, it produced minimal complications while training as both models achieved high accuracy and low loss rates. We theorize that this is likely due to the advance classification methods of both algorithms, but it is difficult to confirm this theory, making it a reasonable subject for further research. Due to the space constraints present in this investigation, it is impractical to review this topic in detail. This is why the data imbalance was not a significant concern for the study, as it likely is due to the advanced feature extraction methods used by both CNNs.

Furthermore, the dataset we used for this study contained images that varied slightly in presentation. Although the images were of high quality, the variation in the types of the pictures may have hindered with the algorithm's ability to effectively learn the dataset. For example, the dataset manufacturers captured images from various angles, including front, back, and top views; despite this limitation, the algorithms attained high accuracies. However, the absence of this limitation could have led to even higher accuracy levels.

To further maximize accuracy, more research regarding Alopecia Areata classification should be conducted using TL with complex networks like NASNET-Large. As shown by this study, the deeper the network, the probability of higher accuracy increases. This theory was also confirmed by another paper (20). However, another potential factor that may have led to the high performance with both algorithms is the feature extraction process. Specifically, it may have been easier for the algorithm to effectively classify between the distinct images, due to the unique features in the images. This may have been a potential issue in other networks. Due to the limited capabilities of the device used in this study, it was impossible to train NASNET-Large, otherwise, this network would have been trained as well. The use of NASNET-Large could allow for another useful, and potentially the most accurate algorithm for the dermatological industry.

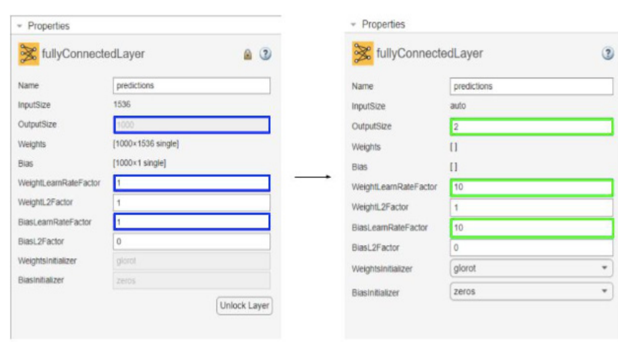

Figure 5: Modifications in the final learnable layer of Inception-Resnet-v2. An example of the applied changes to the final learnable layer of Inception-Resnet-v2. The WeightLearnFactor and BiasLearnFactor were changed to 10, as this would enhance the speed of learning in the modified layer.

https://doi.org/10.59720/24-004

For further research, incorporating an algorithm capable of accurately classifying distinct stages, whether early or late, of Alopecia would be very useful. Since our models were binary, they are unable to report whether someone is in an early stage of Alopecia, which was also a limitation in existing algorithms. Perhaps developing a non-binary algorithm using a similar TL approach could significantly improve diagnostic precision and accuracy. For simplicity's sake, we decided to make a binary model, as we wanted to merely maximize the accuracy of our model and minimize the false positive and false negative rates. However, the addition of this new non-binary algorithm will allow individuals to promptly pinpoint alopecia and receive adequate treatment.

Overall, the results from this study demonstrate the effectiveness and convenience of TL methods not only in dermatological settings, but in the general computer vision field. Modifying a pre-trained neural network to suit a new task is much more efficient than training a network from scratch, and it generally yields higher accuracy rates, which is vital for network optimization. Additionally, this paper offers a new CNN network for Alopecia Areata classification that is relevant in clinical settings.

MATERIALS AND METHODS

We performed algorithm development and analysis using MATLAB, a numerical computing software. We conducted this project in consecutive stages: in summary, the first two involve modifying and training Inception-resnet-v2 and Xception, respectively, and the following step includes evaluating both networks.

For training, we collected a total of 1050 healthy hair images from the Figaro1k dataset (11). For the unhealthy dataset, we extracted a total of 171 relevant Alopecia Areata images from two publicly available datasets: Kaggle, Dermnet, (12) and Scalp Disease detection (5), an independently created dataset. These images conveniently include both the earlier and later stages of Alopecia Areata. To prevent overfitting, we used an 80-20% data split, with 80% of the total data used for training and 20% for validation. This type of data split generally yields the highest performance levels (20).

We used TL to develop the new algorithms. Prior to training, we modified both neural networks in the same way. First, we adjusted the last learnable layer to suit the new task by changing the total number of classes to 2, as the classification was binary. Then, in the same layer, we changed the WeightLearnFactor and BiasLearnFactor to 10 to enhance the speed of learning in the modified layer

Figure 6: Modifications in the final Output Layer of Inception-ResNet-v2 and Xception. An example of the applied changes to the final learnable layer of Inception-ResNet-v2 and Xception. The final Output layer was modified and the Classes and OutputSize categories were set to auto, as these specifications would suit the new task.

Configurations Settings	Selected Options
Optimizer	Stochastic gradient descent with momentum (SGDM)
Momentum	0.9
Initial Learning Rate	0.01
MiniBatchSize	128
MaxEpochs	30
Validation Frequency	50
LearnRateDropFactor	0.1
LearnRateDropPeriod	10
L2Regularization	0.0001

Table 2: Optimization Configurations used for both networks. MATLAB numerical computing software was used to optimize the Convolutional Neural Networks.

(**Figure 5**). Next, we modified the final Output layer by setting the Classes and OutputSize categories to auto, as these specifications would suit the new task. The output layer for both classes was identical (**Figure 6**). Finally, we applied the same Optimization Configurations for both algorithms before algorithm training (**Table 2**).

After the three models finished training, we evaluated their performances in detail. This quantitative assessment assessed the developed frameworks based on 2 relevant criteria. First, we evaluated the trained frameworks based on Criterion 1, which focused on measuring the validation accuracy, and Criterion 2, which focused on evaluating the final validation loss. These criteria served as simple, yet meaningful metrics for assessing accuracy. We defined the model with the higher validation accuracy and lower validation loss rates as the most useful algorithm in clinical settings.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Anuj Tiwari for reviewing this paper and providing relevant, constructive feedback.

Received: December 18, 2023
Accepted: March 18, 2024
Published: June 3, 2024

REFERENCES

1. "Home." *National Alopecia Areata Foundation*, www.naaf.org/. Accessed 1 Dec. 2023.
2. "Alopecia Areata." *Centers for Disease Control and Prevention*, www.niams.nih.gov/health-topics/alopecia-areata#:~:text=Causes%20of%20Alopecia%20Areata,genetic. Accessed 1 Dec. 2023.
3. "Alopecia Areata." *Pfizer*, www.pfizer.com/disease-and-conditions/alopecia-areata#:~:text=of%20Alopecia%20Areata-,An%20estimated%206.7%20million%20people%20in%20the%20United%20States%20and,a%20lifetime%20risk%20of%202%25.&text=That%20estimated%20lifetime%20risk%20for,from%201%25%20to%202%25. Accessed 1 Dec. 2023.

https://doi.org/10.59720/24-004

4. Kim, J., et al. "Hair Follicle Classification and Hair Loss Severity Estimation Using Mask R-CNN." *Journal of Imaging*, vol. 8, no. 10, October 2022, https://doi.org/10.3390/jimaging8100283.
5. Roy, M. "Hair and Scalp Disease Detection using Machine Learning and Image Processing." *European Journal of Information Technologies and Computer Science*, vol. 3, no. 1, 2023, pp. 7–13, https://doi.org/10.24018/compute.2023.3.1.85.
6. Chang, W. et al. "ScalpEye: A Deep Learning-Based Scalp Hair Inspection and Diagnosis System for Scalp Health." *IEEE Access*, vol. 8, July 2020, pp. 3565-3572, https://doi.org/10.1109/ACCESS.2020.3010847.
7. Shih, H. "An Unsupervised Hair Segmentation and Counting System in Microscopy Images." *IEEE Sensors Journal*, vol. 15, no. 6, June 2015, pp. 3565-3572, https://doi.org/10.1109/JSEN.2014.2381363.
8. Benhabiles, H., et al. "Deep Learning based Detection of Hair Loss Levels from Facial Images." *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2019, pp. 1-6, https://doi.org/10.1109/IPTA.2019.8936122.
9. Kim, M., et al. "Deep-Learning-Based Scalp Image Analysis Using Limited Data." *Electronics*, vol. 12, no. 6, Mar. 2023, p. 1380, https://doi.org/10.3390/electronics12061380.
10. Tan, M., et al. "EfficientDet: Scalable and Efficient Object Detection." *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* June 2020, https://doi.org/10.1109/CVPR42600.2020.01079.
11. Svanera, M., et al. "Figaro, hair detection and segmentation in the wild." *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 933-937, https://doi.org/10.47750/pnr.2022.13.S01.02.
12. Dermnet." *Kaggle*, www.kaggle.com/datasets/shubhamgoel27/dermnet. Accessed 22 Dec. 2023.
13. Mahdianpari, M., et al. "Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery." *Remote Sensing*, vol. 10, no. 7, July 2018, p. 1119, https://doi.org/10.3390/rs10071119.
14. Kapoor, I., et al. "Automated Classification Method for Early Diagnosis of Alopecia Using Machine Learning." *Procedia Computer Science*, vol. 132, 2018, pp. 437-443, https://doi.org/10.1016/j.procs.2018.05.157.
15. Szegedy, C., et al. "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning." *arXiv preprint*, 2016, https://doi.org/10.48550/arXiv.1602.07261.
16. Saraswathi, C., & Pushpa, B. "Attention Balanced Multi-Dimension Multi-Task Deep Learning for Alopecia Recognition." *Indian Journal of Science and Technology*, vol. 16, no. 18, 2023, pp. 1365-1373, https://doi.org/10.17485/IJST/v16i18.29.
17. Shabnam, S., et al. "An Analysis of Alopecia Areata Classification Framework for Human Hair Loss Based on VGG-SVM Approach." *Journal of Pharmaceutical Negative Results*, Sept. 2022, pp. 9-15, https://10.47750/pnr.2022.13.S01.02.
18. Saraswathi, C., & Pushpa, B. "AB-MTEDeep Classifier Trained with AAGAN for the Identification and Classification of Alopecia Areata." *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 3, Jun. 2023, pp. 10895–10900, https://doi.org/10.48084/etasr.5852.
19. Abisha, A., & Bharathi, N. "A Hybrid Feature Extraction and Classification using Xception-RF for Multiclass Disease Classification in Plant Leaves." *Applied Artificial Intelligence*, Feb 2023, https://doi.org/10.1080/08839514.2023.2176614.
20. Podlodowski, L., et al. "An ensemble of Deep Convolutional Neural Networks for Marking Hair Follicles on Microscopic Images." In *Proceedings of the Position Papers of the 2018 Federated Conference on Computer Science and Information Systems*, vol. 16, September 2018, https://doi.org/10.15439/2018F389.