International Journal of Current Bio-Medical Engineering (IJCBE)
Volume 4, Issue 1, January – February (2026)

RESEARCH ARTICLE

# SCALP CONVICAP: AN INTELLIGENT CONVINEXT FRAMEWORK FOR HAIR AND SCALP DISEASE DETECTION USING HYBRID CONVNEXT–TRANSFORMER LEARNING WITH GRAPH AND CAPSULE NETWORKS

Bastin Rogers Cross Joseph[1, *] and Arthi Pandiyan[2]

[1]Research Scholar, Department of School of Computer Science and Technology, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, 641114, India.
[2]Assistant Professor, Department of Electronics and Communication Engineering, GRT Institute of Engineering and Technology, Tiruttani, Tamil Nadu 631209, India.

*Corresponding e-mail: bastinrogers@karunya.edu.in*

**Abstract – The determination of hair and scalp diseases is a challenging dermotological activity which requires accurate and automated image-based analysis to make early diagnosis and to treat the patient in the right way. Conventional manual diagnosis is lengthy, saubjective and limited by the fact that it is restricted by large amounts of scalp image. This study aims art resolving these problems by introducing a new scalp. Deep learning system ConViCap that is a blend of ConvNeXt and Vision Transformer (ConViNeXt) in order to collect local texture details as well as global contextual information. The framework uses BiGaBor pre-processing, graph-based feature refinement (GNN), and Capsule Network classification (CapsNet). For the experimental analysis, the proposed model attains an overall accuracy of 99.47% and an F1 score of 95.47% for efficient hair and scalp disease detection. The proposed Scalp ConViCap model enhances overall accuracy by 7.51%,21.28%' 8.21%, and 9.62% compared to Xception with ReLU, ViT-B/16, EfficientNet-based CNN, and R-CNN with CLAHE, respectively.**

***Keywords*** *– ConvNeXt, Vision Transformer, BiGaBor, Graph-based feature refinement (GNN), Capsule Network, Hair and Scalp Disease.*

## 1. INTRODUCTION

Hair, which is primarily made up of keratin protein, is connected with beauty and masculinity; A person's body contains roughly five million hair follicles, and scalp hair plays an important role in regulating body temperature and insulating the brain from external heat [1]. Beyond its biological function, hair plays an important role in personal identification, cultural expression, self-confidence, and social relationships; it is nevertheless susceptible to a variety of illnesses that impair its structure and health [2]. Poor care, environmental exposure, genetics, stress, and nutritional imbalance can all cause hair damage and problems, including alopecia, telogen effluvium, and scalp psoriasis [3]. Dermatitis has several clinical forms that are influenced by environmental or inherited causes, and fungal infections are common and difficult to treat when pathogenic fungi penetrate the body and cause reinfection [4]. Scalp symptoms such as dandruff and pruritus are prevalent and result from a variety of underlying inflammatory disorders with distinct pathophysiology; nevertheless, systematic and scalp-specific diagnostic approaches are still missing, restricting correct classification and individualised treatment [5].

The lack of a reliable, early, and automated approach for detecting hair and scalp illnesses, due to delayed manual diagnosis, high image variability, and limited standardised datasets, continues to be a key barrier in dermatological healthcare [6]. The study addresses the need for an accurate, efficient, and automated scalp inspection system to replace subjective, time-consuming manual diagnosis and limited access to dermatology professionals [7]. The problem addressed is the lack of a reliable and automated machine learning-based approach for correct diagnosis of alopecia areata, resulting in subjective assessment and delayed hair loss identification [8]. Accurate analysis of dermoscopic skin lesion pictures is hampered by hair occlusion, making effective hair recognition and segmentation a significant problem for automated diagnosis [9]. Although the advancement is fast, the AI hair analysis methods require standardization, clinical validation, and simple integration in regular dermatological treatment.

The article addresses the issue of machine learning-based seizure detection transfer algorithm in trustworthy, clinically meaningful real-time scalp electroencephalography (EEG) [11]. This work is meant to do an in-depth surveillance of recent skin image processing techniques in the example of automated diagnosis and to determine current, quantitative evaluation of dermatitis, vitiligo, and alopecia areata, required capabilities, limits, and gaps in research of clinical decision support systems [12]. The paper seeks to discuss and analyse how artificial intelligence is utilised to diagnose, classify, locate, and assess the severity of skin problems in aesthetic medicine, highlighting current obstacles and research needs [13]. The current method combines dermoscopic image-based hair cluster detection models with deep learning techniques to automatically identify and analyse hair clusters for scalp and hair condition assessment [14]. It depends on standard image processing techniques paired with classical machine learning classifiers such as SVM and KNN for detecting and categorising hair and scalp illnesses [15]. Existing research is primarily based on standalone CNN or transformer models, which limit their ability to capture fine-grained local textures and long-range contextual correlations in hair and scalp illness images.

- The proposed Scalp ConViCap model that hybridises ConViNeXt by enabling feature extraction by enhancing the accuracy of hair and scalp disease detection.
- BiGabor Filter is used for enhancing the quality of the Images through minimising the noise and upgrading the image so that the disease can be detected easily.
- The extracted features are refined using a Graph Neural Network (GNN) for robust and reliable disease detection.
- The refined image is classified using the CapsNet classifier, and it classifies the disease to Normal scalp, Alopecia Areata, Seborrheic Dermatitis, Psoriasis, and Tinea Capitis.

## 2. LITERATURE SURVEY

Hair and scalp illnesses are frequent dermatological conditions that have a substantial impact on someone's appearance, psychological well-being, and quality of life. Traditional diagnosis is time-consuming and subjective due to its reliance on visual inspection and clinical knowledge. Automated hair and other technologies enable AI and image processing to progress, which will allow automated hair identification. The identification of scalp disease has become popular in recent years. Machine learning and deep learning techniques that are based on learning have shown promising outcomes in terms of diagnostic precision, reliability and timely identification of scalp diseases.

In 2023, Mrinmoy Roy and Anica Tasnim Protity., [16] offer a 2D Convolutional Neural Image preprocessing method, like data balancing, to categorize hair loss and scalp-related disorders. The training accuracy of the suggested approach was 96.2%, and a validation accuracy of 91.1%, suggesting positive results in the detection of alopecia, psoriasis, and folliculitis.

In 2024, Mohammad Sayem Chowdhury et al., [17] utilizes a Xception deep learning model modified with ReLU activation, dense layer, global pooling regularization and dropout. The model achieved a classification accuracy of 92%, outperforming current architectures like VGG19, Inception, ResNet and DenseNet.

In 2023, Mr Hrishabh Tiwari et al., [18] suggested methods that incorporate deep learning-based scalp diagnosis systems based on CNN and FCN, and machine learning classifiers similar to SVM and KNN to identify disorders of the scalp. These methods performed well, with accuracies and average precision ranging from 82.9% to 99.09%, including up to 97.41% to 99.09% precision and 91.4% accuracy in identifying healthy scalp and alopecia conditions.

In 2024, Changjin Ha et al., [19] present an intelligent healthcare platform for automated scalp condition diagnosis that incorporates deep learning models such as ResNet-152, EfficientNet-B6, and Vision Transformer (ViT-B/16), as well as an attention rollout technique for explainable decision-making. Among the models tested, the ViT-B/16 performed the best, with an average classification accuracy of 78.31% across six scalp hair condition severity categories.

In 2022, Jeong, J.I., et al., [20] proposed an AI-based intelligent scalp diagnostic and classification system called AI-ScalpGrader, which uses a portable scalp imaging device and a deep-learning EfficientNet-based CNN model to identify and categorise 10 different scalp disorders. The suggested method showed high diagnostic reliability, with accuracies ranging from 87.3% to 91.3% in identifying scalp symptoms and severity levels.

In 2023, Hansoo Kim., [21] suggested a technique that provides a novel deep learning-based system for rapid and automatic scalp condition classification based solely on scalp images, with no need for additional user metadata. The method had a maximum accuracy of 98.78% and an average accuracy improvement of 6.62% over traditional approaches.

In 2024, Honghai Ji., et al., [22] came up with a multi-network fusion object detection framework that is a combination of Faster R-CNN algorithm and contrast-limited adaptive histogram equalisation (CLAHE) to enhance dermoscopic images to the automated diagnosis of scalp psoriasis and seborrheic dermatitis. The proposed approach achieved 91.0% diagnostic accuracy for scalp psoriasis and 89.9% accuracy in identifying key dermoscopic patterns, placing it on par with dermatologists' performance.

In 2023, C Saraswathi and B Pushpa., [23] offers an Attention-based Balanced Multi-Task Deep (AB-MTDeep) learning framework that combines Multi-Task Learning and Cross-Residual Learning to learn local and global characteristics from hair and scalp photos for Alopecia Areata classification. The proposed system had better performance over the traditional methods, and an accuracy of classification that is 95.11%.

In 2022, Mrinmoy Roy and Anica Tasnim Protity., [24] proposed a method of hair and scalp disease detection that is a hybrid feature extraction method that uses image processing and general machine learning classifiers, mostly SVM and KNN. The model achieved a general classification accuracy of about 90% with good performance in the detection of scalp-related diseases.

In 2024, Nair, S.A.L. and Megalingam, R.K., [25] determine human attention levels using neural signals, the proposed technique combines deep learning models and a brain-computer interface (BCI). The system achieved an attention detection accuracy of more than 90%, confirming the efficiency of the deep learning-BCI integration.

## 3. SCALP CONVICAP METHODOLOGY

The proposed Scalp ConViCap architecture begins with the input scxalp images, which are enhanced with the help of a specific BiGabor preprocessing filter to eliuminate noise without losing texture and lesion edges. The data augmentation is then applied on the already processed images to enhance resilience and generalisation. ConViNeXt is a deep discriminating feature extractor that extracts local and global features of the scalp. Such properties are additionally augmented with Graph-based Neural Network (GNN), a neural unit that simulates the interaction between features and enhances disease-specific representations. The last is a Capsule Network (CapsNet) is used to classify normal scalp and disease affected scalp and to quantitatively assess the system performance in terms of accuracy, precision, specificity, recall, and F1 score Figure 1.

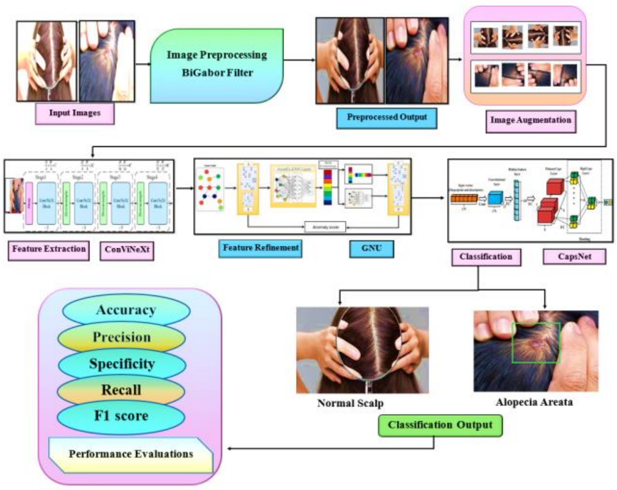



**Figure 1:** Architecture of a Scalp ConViCap Methodology

### *3.1 DermNet Scalp Hair Dataset*

The DermNet Scalp & Hair Dataset is a free dermatological image dataset of image with free access of more than 1500 clinical images of scalp and hair issues. It contains a diversity of problems, such as alopecia areata, scalp psoriasis, seborrheic dermatitis, tinea capitus, folliculitis and other inflammatory or infectious scalp diseases. The images were captured in actual clinical conditions with significant variations in light intensity, hair density, lesions and disease severity. All images are curated and identified by dermatologists to make them clinically credible. This dataset is often used to train and test machine learning and deep learning models on scalp and hair diseases, automated detection and classification.

### *3.2 BiGabor Filter*

The BiGabor filter is a new hybrid pre-processing method that is a combination of the advantages of the Bilateral and Gabor filters in enhancing the shapes of lesions of the scalp and preserving fine hair textures. Hair needs noise reduction, edge preservation, as well as texture enhancement and scalp disease images are caused because of the existence of thick hair stands, uneven light, and covert pathological patterns. The proposed BiGabor filter addresses these problems by reducing the noise without losing the boundaries of the lesion, and then points out directional and frequency-based texture characteristics of relevance to the diseases of the scalp.

The Bilateral Filter achieves edge-preserving smoothing by considering both proximity of the space and similarity of intensity. The mathematical equation is as follows:

$$I_{BF}(m) = \frac{1}{P(m)} \sum_{n \epsilon \Omega} I(n) exp\left(-\frac{\|m-n\|^2}{2\sigma_q^2}\right) exp\left(-\frac{|I(m)-I(n)|^2}{2\sigma_t^2}\right) \quad (1)$$

where $I(m)$ and $I(n)$indicate pixel intensities, $\sigma_q$ regulates and gives control over spatial smoothing, $\sigma_t$ controls intensity similarity, and $P(m)$ is the normalization factor. This process is effective in the removal of preservation artifacts and noise during acquisition and retention of scalp lesions and hair boundaries.

After bilateral smoothing, the Gabor Filter is applied to obtain orientation and frequency- specific texture information on hair shafts, follicular patterns, scaling and lesion borders. A Gabor filter is a two-dimensional filter that is described as follows:

$$R(m,n) = exp\left(-\frac{m'^2+\gamma^2 n'^2}{2\sigma^2}\right) cos\left(2\pi \frac{m'}{\lambda} + \emptyset\right) \quad (2)$$

Where $m' = m\, cos\theta,\ n' = -m\, sin\theta + n\, cos\theta$

In this case, $\lambda$ is the wavelength, $\theta$ is the direction, $\sigma$ is the Gaussian envelope width, $\gamma$is the aspect ratio, and $\phi$ is the phase offset. This step enhances disease-specific textural indicators such as patchy hair loss, scaling patterns, and inflamed areas.

The two operations, when combined one after the other, produce the final BiGaBor output:

$$I_{BiGabor} = R(I_{BF}) \quad (3)$$

This hybrid filtering technique enhances features and reduces spurious texture noise because of hair boundary, and enhances lesion contrast. Consequently, the BiGabor filtered images are better input to deep learning models, which leads to increased accuracy in hair and scalp disease detection and classification.

### *3.3 ConViNeXt*

The proposed Scalp ConViCap architecture integrates the high local feature extracting capacity of ConvNeXt and the global contextual modelling capacity of the Vision Transformer to facilitate effective representation learning to identify hair and scalp diseases. The approach uses concurrent CNN- and Transformer-based branches, followed by successful feature fusion, to generate a discriminative hybrid feature representation.

The ConvNeXt branch specialises in extracting fine-grained local features such as hair follicles, textural irregularities, lesion borders, and scalp surface variations. Each ConvNeXt block uses depth-wise convolution followed by pointwise convolution to improve spatial feature learning while minimising computing expense.

The convolutional feature extraction is described as:

$$R_h = \sigma\left(PWConv\left(DWConv(J)\right)\right) \tag{4}$$

The notation $DWConv(\cdot)$ refers to depthwise convolution, $PWConv(J)$ to pointwise convolution, and $\sigma(\cdot)$ to nonlinear activation. Hierarchical representations are learned using several stacked ConvNeXt layers as follows:

$$S_h^m = f_h^m(S_h^{m-1}), \qquad m = 1,2,3,\dots M \tag{5}$$

The final result $S_h$ collects detailed spatial and structural information needed to detect localized scalp anomalies.

The Vision Transformer branch is used to model long-term dependencies and capture global contextual linkages throughout the scalp area. The input image is initially separated into non-overlapping patches, which are linearly projected into embedding vectors.

$$N_k = A_k E_c + a_c \tag{6}$$

Where $A_k$ represents the $i$th picture patch and $E_c, a_c$ are learnable parameters. The embedded patches are processed using multi-head self-attention, defined as follows:

$$Attention(X, B, W) = softmax\left(\frac{XB^Q}{\sqrt{d_g}}\right)V \tag{7}$$

This approach allows the model to learn global relationships among distant scalp regions. The transformer encoder output is expressed as follows:

$$S_u = Transformer\ Encoder(N) \tag{8}$$

To capitalize on the complementary characteristics of both branches, convolutional and transformer features are combined via a feature fusion layer:

$$F_{fusion} = \phi(S_h \oplus S_u) \tag{9}$$

The symbol $\oplus$ represents feature concatenation, while $\phi(\cdot)$ indicates normalization and dimensional alignment. The combined features are also enhanced in order to achieve the final hybrid representation:

$$R_{hybride} = v\left(R_{fusion}\right) \tag{10}$$

The hybrid feature vector generated is the combination of local texture feature with global semantic context, and thus it is so helpful in detecting hair and scalp diseases.

The proposed ConViNeXt architecture enhances feature representation, generalisation, and performance of the hair and scalp disease detection method by integrating the ConvNeXt and Vision Transformer architectures in Figure 2.

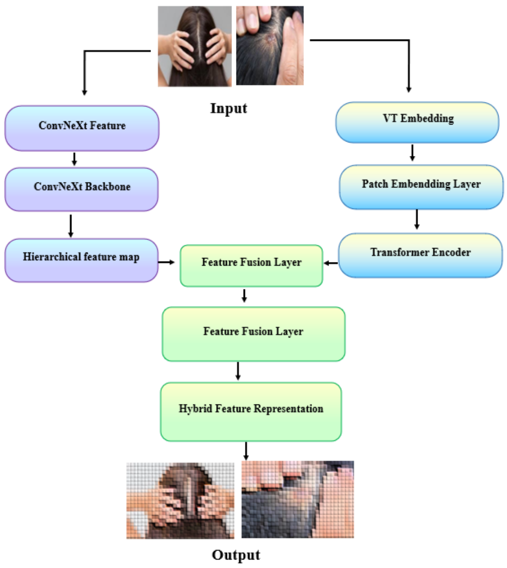



**Figure 2:** ConViNeXt Architecture

### *3.4 GNN*

Duplicated information is often found in fused deep features of hair and scalp disease detection because of hair density differences, lighting and background scalp texture. To overcome this problem, we adopt a Graph Neural Network (GNN)-based feature refinement block to enhance the disease-related patterns by directly modelling the connections among the areas of the scalp. Compared to the conventional attention operations, GNN-refinement identifies structural, geographical, and semantic links between places like follicles, lesions, inflamed regions and adjacent skin.

Let the fused feature map acquired from previous phases be denoted as follows:

$$R_{fusion} \in \mathbb{R}^{S\times E\times D} \tag{11}$$

The feature map is divided into N meaningful sections (patches or superpixels), with each represented as a graph node. The resulting graph is defined as follows:

$$F = (\vartheta, \varepsilon) \tag{12}$$

where $\vartheta = \{v_1, v_{2,\dots,} v_n\}$represents nodes and $\varepsilon$ represents edges that encode geographic or similarity-based linkages. Each node $v_i$ is initialized with a feature vector.

$$g_i^{(0)} \in \mathbb{R}^D \tag{13}$$

Extracted from the relevant section of $R_{fusion}$.

The refinement procedure is carried out iteratively, allowing each scalp region to include contextual information from nearby parts. In the $m$-th GNN layer, node features are updated as follows:

$$g_i^{(m+1)} = \sigma\left(\sum_{e\in\mathcal{N}(i)} E^{(m)} g_e^{(m)} + a^{(m)}\right) \quad (14)$$

$\mathcal{N}(i)$denotes the neighbors of node $i$, $E^{(m)}$and $a^{(m)}$are learnable parameters, and $\sigma(\cdot)$represents a nonlinear activation function. This technique improves node characteristics by highlighting consistent disease patterns across linked scalp regions.

An attention mechanism is included in the GNN to draw attention to aberrant scalp regions. The attention coefficient between connected nodes $k$ and $e$ is calculated as follows:

$$\alpha_{ke} = \frac{exp\left(LeakyRELU\left(l^{Q}[Eg_k \| Eg_k]\right)\right)}{\sum_{g\in\mathcal{N}(k)} exp(\cdot)} \quad (15)$$

The refined node update changes to:

$$g_k^{m+1} = \sigma\left(\sum_{e\in\mathcal{N}(k)} \alpha_{ke} E g_e^{(a)}\right) \quad (16)$$

This process prioritizes nodes reflecting sick scalp regions over irrelevant or healthy background areas.

Following the $M$ GNN layers, the refined node features are marked as:

$$g_k^{(M)}, \qquad k = 1,2,\dots,N \quad (17)$$

These revised node representations are combined to reconstitute the refined feature map.

$$R_{ref} = Reconstruct\left(\left\{b_k^{(M)}\right\}\right) \quad (18)$$

Pooling can be used to obtain a global refined feature vector.

$$f_{ref} = GAP\left(R_{ref}\right) \quad (19)$$

Pooling can be used to obtain a global refined feature vector.

The GNN-based feature refinement module improves disease-specific patterns by using relational information from scalp regions. By modelling inter-region interdependence, the approach reduces noise, enhances feature discriminability, and increases robustness to visual fluctuation, resulting in better hair and scalp disease identification.

### *3.5 CapsNet*

To accurately classify hair and scalp diseases, spatial and hierarchical linkages between features such hair follicles, lesions, inflammation regions, and texture patterns must be preserved. Conventional classifiers frequently lose this information as a result of pooling. To address this limitation, a Capsule Network (CapsNet)-based classifier is used to characterize both the presence and spatial configuration of disease-related characteristics, allowing for Figure 3 accurate differentiation of visually comparable scalp diseases.

The refined feature representation derived from the previous feature extraction and refinement phases is moulded into a set of capsules:

$$Z = \{z_1, z_2, \dots, z_N\}, \qquad z_i \in \mathbb{R}^d \quad (20)$$

Each capsule output vector encodes instantiation parameters for scalp features, and its magnitude denotes the likelihood of feature occurrence. A squashing function is used to normalise capsule outputs:

$$\hat{t}_{e|k} = E_{ke} t_k \quad (21)$$

The routing coefficients are generated via softmax normalisation:

$$h_{ke} = \frac{exp(a_{ke})}{\sum_g exp(a_{kg})} \quad (22)$$

The total input to a higher-level capsule is:

$$l_e = \sum_k h_{ke} \hat{t}_{e|k} \quad (23)$$

After squashing, the final capsule output is $w_e$. The length of the output capsule vector shows the risk of a certain hair or scalp disease class:

$$\hat{y}_e = \|w_e\| \quad (24)$$

This is projected onto the longest reaching capsule class to come to the result. The Capsule Network Classifier enhances the resistance and classification accuracy in the detection of hair and scalp disease by using spatial hierarchies and pose information.

The graphical representation shows an entire pipeline of hair and scalp disease detection, starting with the BiGabor-based preprocessing to enhance the texture and directional characteristics of scalp images. This is then followed by data augmentation to enhance image diversity and model generalization and feature extraction with the ConViNeXt architecture to obtain both local and global representation. The obtained features are further refined through a GNN-based feature refinement module, simulating interaction between the regions of the scalp, then generated with a CapsNet classifier to generate disease-specific responses. Lastly, accuracy, precision, recall, specificity and F1 score are used to assess the effectiveness of the proposed model since they give a comprehensive measure of the classification effectiveness.

## 4. RESULT AND DISCUSSION

The experimental implementation of the proposed Scalp ConViCap hair and scalp disease identification model was doneona curated collection of 1500 scalp images, including both normal and pathological conditions, like Alopecia Areata, Dandruff, and psoriasis. A reasonable division of the dataset was made into training, validation and testing sets to allow the evaluation of performance without bias. Each of the experiments was conducted in Python using TensorFlow and PyTorch deep learning libraries, and image processing by OpenCV. The given noel was trained and maintained to work on a platform which involved an Intel Core i7 processor, [16/32] GB RAM, and NVIDIA graphics card running Windows/Linux. Such an arrangement enabled quick learning and reliable assessment of the proposed model.

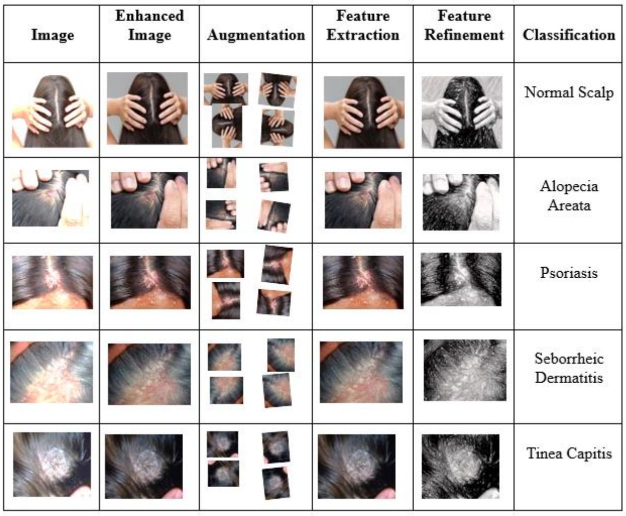



**Figure 3:** Experience outcomes of the Scalp ConViCap model

### *4.1 Performance Analysis*

The Accuracy, precision, recall, specificity, and F1-score are among the common evaluation metrics used to assess the effectiveness of the suggested hair and scalp illness detection method. The results demonstrate that the Scalp ConViCap system is a reliable method in detecting heaslthy and pathological scalp conditions as the classification accuracy is high, with a balanced precision and recall rate. in every way, the performance demonstrates the reliability and generalization of the proposed model on a scale of hair and scalp images.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{25}$$

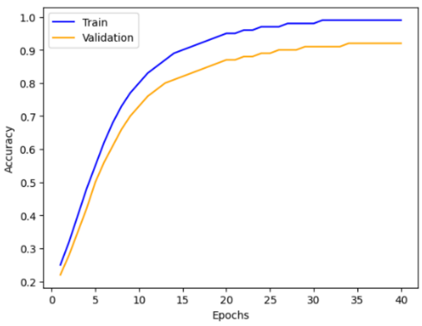
$$Precision = \frac{TP}{TP+FP} \tag{26}$$

$$Recall = \frac{TP}{TP+FN} \tag{27}$$

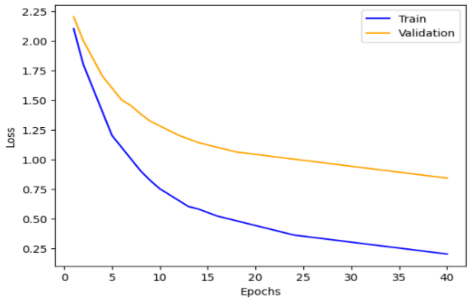
$$Specificity = \frac{TN}{TN+FP} \tag{28}$$

$$F1\ score = 2 * \frac{Precision \times Recall}{Precision+Recall} \tag{29}$$

Where TP, TN, FP, and FN denote True Positive, True Negative, False Positive, and False Negative, respectively.




**(a)**




**(b)**

**Figure 4:** (a) Accuracy curve of the Scalp ConViCap model (b) Loss curve of the Scalp ConViCap model

The training and validation graphs show a consistent and effective learning process for the model throughout epochs. Training accuracy rises Figure 4 (a) quickly in the early stages and progressively converges, but validation accuracy follows a similar pattern with a small gap, indicating good generalisation and minimal overfitting. Similarly, both training and validation losses steadily decrease, indicating consistent Figure 4 (b) optimisation and convergence. Overall, the smooth and well-matched curves demonstrate that the model learns essential characteristics and performs consistently on unseen data.

### *4.2 Comparative Analysis*

The comparison research demonstrates that standalone models like ConvNeXt and DINO-ViT perform moderately, indicating limited capacity to capture both local and global discriminative features. The CMM Ensemble and ConvNeXt+CBAM models greatly increase accuracy and class-wise metrics by utilizing feature fusion and attention techniques. However, the suggested ConViNeXt model outperforms all evaluation criteria, with near-perfect accuracy and significant improvements in precision, recall, specificity, and the F1-score. This demonstrates that the hybrid combination of ConvNeXt and Vision Transformer is a powerful tool to capture both fine-grained texture details and long-range contextual dependencies to generate better and improved results.

**Table 1:** Comparison between DL Networks

| Networks | Accuracy | Precision | Recall | Specificity | F1 score |
|---|---|---|---|---|---|
| ConvNeXt | 69.23% | 65.42% | 62.78% | 68.11% | 64.45% |
| DINO-ViT | 71.65% | 69.24% | 67.62% | 70.99% | 68.75% |
| CMM Ensemble | 83.17% | 72.59% | 70.96% | 77.49% | 82.27% |
| ConvNeXt+CBAM | 85.23% | 76.54% | 74.22% | 81.37% | 85.07% |
| ConViNeXt | 99.47% | 90.82% | 89.12% | 91.42% | 95.47% |

Figure 5 shows a graphical depiction of the comparative performance described in Table 1. The bar chart compares assessment metrics across several deep learning architectures, allowing for a more intuitive understanding of model performance variations. This visualisation demonstrates the proposed ConViNeXt network's consistency and excellence across all measures when compared to existing techniques.

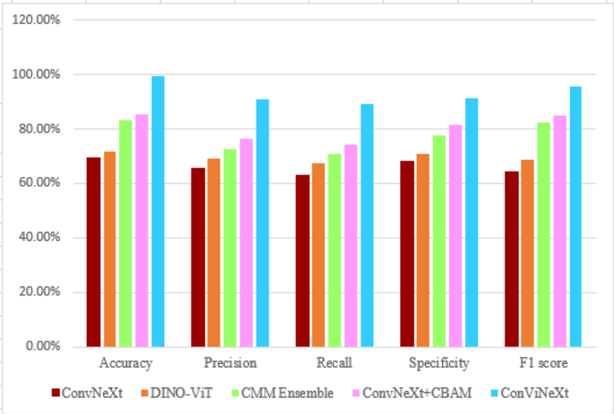



**Figure 5:** Performance Comparison of DL Networks

The bar graph is used to compare results of various deep learning models in terms of accuracy, precision, recall, and specificity and F1 score. Traditional designs, which include ConvNeXt and DINO-ViT, yield mediocre performances, but CMM Ensemble and ConvNeXt+CBAM show considerable gains after incorporating fusion of features and attention. The ConViNeXt proposed model constantly outperforms every other approach on all the metrics, with high results by a significant factor. This illustrates the effectiveness of the ConvNeXt with a Vision Transformer in the ability to combine the local texture features and the global contextual features in order to be accurately categorised

**Table 2:** Accuracy Comparison of existing and proposed model

| Author | Approaches | Accuracy |
|---|---|---|
| Chowdhury, M.S., et al.,[17] | Xception With ReLU | 92 % |
| Changjin Ha et al., [19] | ViT-B/16 | 78% |
| Jeong, J.I., et al., [20] | EfficientNet-based CNN | 91.3% |
| Honghai Ji., et al., [22] | R-CNN with CLAHE | 89.9% |
| PROPOSED | ConViNeXt | 99.47% |

Table 2 describes the experimental settings and involved test samples that were collected to test the suitability of the former frameworks. The overall accuracy of the proposed system is greater than that of the Xception with ReLU, ViT-B/16, EfficientNet-based CNN, and R-based CNN. By 7.51, 21.28, 8.21 and 9.62, respectively, CNN with CLAHE methods were improved.

## 5. CONCLUSION

In this study, a new hybrid deep learning version called Scalp ConViCap was proposed that can accurately detect hair and scalp diseases and appropriately combines ConvNeXt and Vision Transformer, using local texture and global contextual data. The proposed method that succeeded BiGaBor preprocessing, graph-based feature refinement, and Capsule Network classification showed good results on the DermNet scalp dataset, and the results were high. precision and sound generalisation. Although the model is effective, its performance can be subject to small data size, imbalance, and variance in illumination and capture setting of images. The outcome of the experiments showed that the methods proposed achieve a high level of accuracy of diagnosis of 99.47% and an F1 score of 99.47, which implies its reliability and detection. The approach surpasses the current methods by 7.51 percent against Xception. and with ReLU, 21.28% compared to ViT-B/16, 8.21% compared to EfficientNet-based CNN and 9.62% compared to R-CNN with CLAHE. Moreover, the hybrid architecture has the potential to be too computationally intense to run on low-resource computers. Future work will concentrate on large-scale clinical validation, improving model efficiency, expanding the framework to multi-class disease severity grading, and incorporating explainable AI methodologies to aid dermatological decision-making and real-world clinical adoption.

## CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

## FUNDING STATEMENT

Authors did not receive any funding.

## ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

## REFERENCES

[1] M. Roy and A.T. Protity, "Hair and scalp disease detection using machine learning and image processing," arXiv Preprint arXiv:2301.00122, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] D.N.N. Anh, P.N. Giau, B.N. Le Nguyen, A. Mai, N.T. Hien, and N.T.V. Tuyen, "Harnessing deep learning for scalp and hair disease classification: A comparative study of convolutional neural networks architectures," Journal of Digital Information Management, vol. 23, no. 2, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[3] I.M.K. Karo, D. Kiswanto, S. Panggabean, and A. Perdana, "Hair Disease Classification Using Convolutional Neural Network (CNN) Algorithm with VGG-16 Architecture," Sinkron: Jurnal dan Penelitian Teknik Informatika, vol. 7, no. 4, pp. 2786–2793, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] T. Alyas, K. Alissa, A.S. Mohammad, S. Asif, T. Faiz, and G. Ahmed, "Innovative Fungal Disease Diagnosis System Using Convolutional Neural Network," Computers, Materials & Continua, vol. 73, no. 3, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] B.R. Kim, S.H. Won, J.W. Kim, M. Kim, J.I. Jeong, J.W. Shin, C.H. Huh, and J.I. Na, "Development of a new classification and scoring system for scalp conditions: Scalp Photographic Index (SPI)," Journal of Dermatological Treatment, vol. 34, no. 1, p. 2181655, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] K. Sultanpure, B. Shirsath, B. Bhande, H. Sawai, S. Gawade, and S. Samgir, "Hair and scalp disease detection using deep learning," arXiv Preprint arXiv:2403.07940, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] S.Y. Jhong, P.Y. Yang, and C.H. Hsia, "An Expert Smart Scalp Inspection System Using Deep Learning," Sensors & Materials, vol. 34, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] S. Sayyad and D. Midhunchakkaravarthy, "Deep study on alopecia areata diagnosis for hair loss related autoimmune disorder problem using machine learning techniques," Available at SSRN 4054490, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] N. Lama, R. Kasmi, J.R. Hagerty, R.J. Stanley, R. Young, J. Miinch, J. Nepal, A. Nambisan, and W.V. Stoecker, "ChimeraNet: U-Net for hair detection in dermoscopic skin lesion images," Journal of Digital Imaging, vol. 36, no. 2, pp. 526–535, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] T. Nawaz, G. Williams, and J. Yoo, "Exploring the Advancements and Applications of AI-Powered Hair Analysis Tools," J Clin Aesthet Dermatol, vol. 19, no. 2, pp. 58–62, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[11] N. Moutonnet, S. White, B.P. Campbell, S. Sanei, T. Tanaka, H. Ji, D. Mandic, and G. Scott, "Clinical translation of machine learning algorithms for seizure detection in scalp electroencephalography: systematic review," arXiv Preprint arXiv:2404.15332, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] A. Kallipolitis, K. Moutselos, A. Zafeiriou, S. Andreadis, A. Matonaki, T.G. Stavropoulos, and I. Maglogiannis, "Skin image analysis for detection and quantitative assessment of dermatitis, vitiligo and alopecia areata lesions: a systematic literature review," BMC Medical Informatics and Decision Making, vol. 25, no. 1, p. 10, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[13] A.K.W. Lee, L.K.W. Chan, C.H. Lee, J.M.C. Bohórquez, D. Haykal, J. Wan, and K.-H. Yi, "Artificial intelligence application in diagnosing, classifying, localizing, detecting and estimating the severity of skin conditions in aesthetic medicine: A review," Dermatological Reviews, vol. 6, no. 1, p. e70015, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[14] Y. Xiong, K. Yu, Y. Lan, Z. Lei, and D. Fan, "Hair cluster detection model based on dermoscopic images," Frontiers in Physics, vol. 12, p. 1364372, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] M. Roy and A.T. Protity, "Hair and scalp disease detection using machine learning and image processing," arXiv Preprint arXiv:2301.00122, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] M. Roy and A.T. Protity, "Hair and scalp disease detection using machine learning and image processing," arXiv Preprint arXiv:2301.00122, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] M.S. Chowdhury, T. Sultan, N. Jahan, M.F. Mridha, M. Safran, S. Alfarhood, and D. Che, "Leveraging deep neural networks to uncover unprecedented levels of precision in the diagnosis of hair and scalp disorders," Skin Research and Technology, vol. 30, no. 4, p. e13660, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] H. Tiwari, J. Moolchandani, and S. Mantri, "Diagnosis of Scalp Disorders using Machine Learning and Deep Learning Approach--A Review," arXiv Preprint arXiv:2308.07052, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] C. Ha, T. Go, and W. Choi, "Intelligent healthcare platform for diagnosis of scalp and hair disorders," Applied Sciences, vol. 14, no. 5, p. 1734, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] J.I. Jeong, D.S. Park, J.E. Koo, W.S. Song, D.J. Pae, and H.J. Choi, "Artificial intelligence (AI) based system for the diagnosis and classification of scalp health: AI-ScalpGrader," Instrumentation Science & Technology, vol. 51, no. 4, pp. 371–381, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] H. Kim, "Development of scalp diagnosis algorithm using surface-sensing convolutional neural network," National Academy Science Letters, vol. 47, no. 2, pp. 153–156, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] H. Ji, J. Li, X. Zhu, L. Fan, W. Jiang, and Y. Chen, "Enhancing assisted diagnostic accuracy in scalp psoriasis: A Multi-Network Fusion Object Detection Framework for dermoscopic pattern diagnosis," Skin Research and Technology, vol. 30, no. 4, p. e13698, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] C. Saraswathi and B. Pushpa, "Attention Balanced Multi-Dimension Multi-Task Deep Learning for Alopecia Recognition," Indian Journal of Science and Technology, vol. 16, no. 18, pp. 1365–1373, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] M. Roy and A.T. Protity, "Hair and scalp disease detection using machine learning and image processing," arXiv Preprint arXiv:2301.00122, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] S.A.L. Nair and R.K. Megalingam, "Human attention detection system using deep learning and brain–computer interface," Neural Computing and Applications, vol. 36, no. 18, pp. 10927–10940, 2024. [CrossRef] [Google Scholar] [Publisher Link]


**AUTHORS**

**Bastin Rogers Cross Joseph** and received his B.E. degree in Computer Science and Engineering from Anna University, Chennai, India in the year 2007. He received his M.Tech. degree in Computer Science and Engineering from Karunya University, Coimbatore, India in the year 2011. His areas of interest include Medical Image Processing, IoT, and Network Security.

**Arthi Pandiyan** working as an Assistant Professor in the Department of Electronics and Communication Engineering in GRT Institute of Engineering and Technology- Tiruttani. I am having an experience of 15 years in Engineering teaching field. I received my Bachelor degree in the field of Engineering in the stream of ECE from Anna University and M.E with specialization in Applied Electronics from Anna University. I am currently pursuing my Doctorate in Anna University.

Arrived: 15. 01. 2026

Accepted: 16. 02. 2026