Scalp Diagnostic System With Label-Free Segmentation and Training-Free Image Translation

Youngmin Kim[†1], Saejin Kim[†1], Hoyeon Moon[1], Youngjae Yu[‡2], and Junhyug Noh[‡3]

[1] Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul, Korea
[2] Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul, Korea
[3] Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, Korea
mycalljordan@snu.ac.kr, junhyug@ewha.ac.kr

Abstract. Dermatoses of the scalp affect millions of people around the world, underscoring the urgent need for early diagnosis and management of the disease. However, the development of a comprehensive AI-based diagnosis system encompassing these conditions remains an underexplored domain due to the challenges associated with data imbalance and the costly nature of labeling. To address these issues, we propose "ScalpVision", an AI-driven system for the holistic diagnosis of scalp diseases. In ScalpVision, effective hair segmentation is achieved using pseudo image-label pairs and an innovative prompting method in the absence of traditional hair masking labels. Additionally, ScalpVision introduces *DiffuseIT-M*, a generative model adopted for dataset augmentation while maintaining hair information, facilitating improved predictions of scalp disease severity. Our experimental results affirm ScalpVision's efficiency in diagnosing a variety of scalp conditions, showcasing its potential as a valuable tool in dermatological care. Our code is available at https://github.com/winston1214/ScalpVision.

Keywords: Scalp Disease Diagnosis · Generative Data Augmentation.

1 Introduction

Scalp disorders are a widespread concern, with nearly 90% of adults in the U.S. experiencing some form of condition [6]. Left unchecked, even seemingly mild scalp ailments can escalate into more serious outcomes, such as alopecia, underscoring the importance of timely intervention. Consequently, early diagnosis is crucial for preventing the progression of scalp-related diseases [16, 17], highlighting the need for advanced diagnostic approaches that are both efficient and accessible. Recognizing the importance of early detection, numerous studies have explored scalp disease diagnosis using microscopic scalp imagery [4, 11, 19].

Nevertheless, effectively diagnosing scalp disorders relies heavily on measuring critical features such as hair count and thickness, which demands precise

[†] Equal contribution [‡] Co-supervision

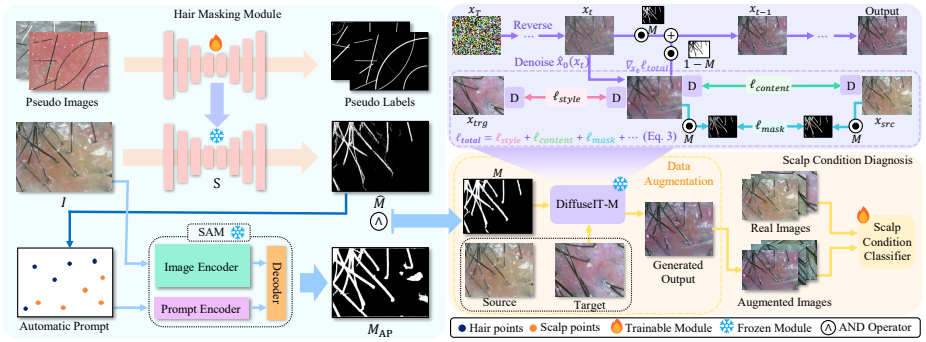

Fig. 1: ScalpVision pipeline overview: I is the original image, model S generates the hair segmentation mask $\hat{M}$ using a pseudo-training set, M_{AP} is the SAM-produced mask, and M is the combined hair segmentation mask. The "Automatic Prompt" for refining segmentation comes from $\hat{M}$. x_{src} and x_{trg} are the source and target images, with M as the mask image of x_{src}. The weighted image sum is denoted by $\odot$ and D stands for DINO-ViT [3].

hair segmentation. However, generating pixel-level hair annotations is costly and time-consuming, and no publicly available dataset provides such segmentation labels. The only major resource, AI-Hub [1], offers classification labels for scalp conditions but lacks segmentation annotations (see Section 3.1). Moreover, like many scalp image datasets, it suffers from data imbalance, especially for severe conditions, making it challenging to develop robust models.

To overcome these limitations, we propose ScalpVision, a comprehensive system for the in-depth assessment of scalp health. First, we achieve label-free hair segmentation by combining a naive segmentation model – trained on synthetic image-label pairs – with an *automatic prompting* module for the Segment Anything Model (SAM) [12], systematically generating positive and negative point prompts to enable accurate hair masks without manual labeling. Building on these masks, we then introduce *DiffuseIT-M*, a diffusion-based image-to-image translation framework that preserves hair details while altering scalp conditions. By generating diverse training samples, our method effectively mitigates data imbalance, ultimately leading to enhanced diagnostic performance for scalp diseases.

2 Method

As illustrated in Figure 1, central to ScalpVision is a hair segmentation module (Section 2.1) and an image translation module for generating diverse scalp images to augment training datasets for scalp condition classification (Section 2.2).

Algorithm 1: Extraction of representative points from mask

Input: Mask $\hat{M}$, bounding box size n, cross-shaped structuring element *kernel*
Output: Representative hair points from mask $\hat{\mathcal{C}}$

$\mathcal{H}_{\text{copy}} \leftarrow \hat{M}$; $\hat{\mathcal{H}}_{\text{skel}} \leftarrow$ zero array with same size as $\mathcal{H}_{\text{copy}}$; $\hat{B}, \hat{\mathcal{C}} \leftarrow \{\}$
while $\mathcal{H}_{copy} \neq 0$ **do**
  Eroded $\leftarrow$ `MORPHOLOGY_ERODE`($\mathcal{H}_{\text{copy}}$, *kernel*)
  Dilated $\leftarrow$ `MORPHOLOGY_DILATE`(*Eroded*, *kernel*)
  $\hat{K} \leftarrow \mathcal{H}_{\text{copy}} - \textit{Dilated}$; $\hat{\mathcal{H}}_{\text{skel}} \leftarrow \hat{\mathcal{H}}_{\text{skel}} \vee \hat{K}$; $\mathcal{H}_{\text{copy}} \leftarrow \textit{Eroded}$
foreach $(x, y) \in \mathcal{H}_{skel}$ **do**
  $\hat{B} \leftarrow \hat{B} \cup \{(x - \frac{1}{2}n, y - \frac{1}{2}n, x + \frac{1}{2}n, y + \frac{1}{2}n)\}$
$\hat{B} \leftarrow$ `NMS`($\hat{B}$)
foreach $(x_1, y_1, x_2, y_2) \in \hat{B}$ **do**
  $\hat{\mathcal{C}} \leftarrow \hat{\mathcal{C}} \cup \{(\bar{x}, \bar{y})\}$ as in Eq.(1)
return $\hat{\mathcal{C}}$

2.1 Label-Free Hair Segmentation

For the precise diagnosis of scalp conditions, our initial step involves segmenting hair within microscopic scalp images. However, since most scalp condition datasets lack segmentation labels, supervised learning methods are not feasible.
Heuristic-driven pseudo-labeling. To address the absence of hair segmentation, we first generate pseudo labels for training our segmentation model (S as shown in Figure 1) using prior knowledge. With the intuition that the hair on the microscopic scalp images follows either a linear function or a power function, we generate synthetic images to effectively guide the model to learn hair patterns on the scalp images. For each disease condition, we randomly select one image representing each distinct severity level, extract three smaller patches from regions of the scalp with no visible hair, and draw curves to simulate hair patterns. Additionally, to simulate dandruff noise, circular white shapes are added to these patches but are not indicated in the pseudo masks, thus training the model to interpret them as noise. We generate 3,000 pseudo-images and corresponding pseudo mask labels, using them to train the U^2-Net [18] which generates the mask, $\hat{M} = [\hat{M}(i,j)] \in [0,1]^{H \times W}$, where H and W are the height and width of the image, and $i \in [1, H]$, $j \in [1, W]$ denote pixel coordinates.
Automatic prompting for SAM. To refine the hair segmentation mask $\hat{M}$, we utilize the foundation segmentation model, SAM [12], employing a point-prompting method to differentiate hair from scalp without additional training. However, selecting random points from $\hat{M}$ for positive prompts often led to suboptimal masks, mainly due to points near the edges of $\hat{M}$ confusing the SAM. Furthermore, the intrinsic randomness occasionally caused sampled points to coalesce within a confined region, thereby leading the SAM to segment only a limited subset of hairs. To address these issues, we developed an automatic prompting method, shown in Algorithm 1, that uniformly samples across $\hat{M}$ and uses the coarse segmentation mask $\hat{M}$ to guide the SAM with high confidence.

To extract the distinct features of the hair, we compute the skeletonized mask, $\hat{\mathcal{H}}_{\text{skel}} \in \{0,1\}^{H\times W}$, using morphological erosion and dilation following [23]. Then, we generate bounding boxes around each pixel in $\hat{\mathcal{H}}_{\text{skel}}$ with size $n \times n$ where we set $n = 10$. These boxes undergo non-maximum suppression (NMS) to filter out the bounding boxes, denoted as $\hat{B} = \{\hat{b}_j\}_{j=1}^{k}$, where each box is defined by coordinates $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$. Following this, the mean points of the hair pixels, $\hat{\mathcal{C}} = \{\hat{c}_j\}_{j=1}^{k}$, in each bounding box $\hat{B}$ can be determined. For each $\hat{b}_j = (x_1, y_1, x_2, y_2)$, the mean point $\hat{c}_j = (\bar{x}, \bar{y})$ is given by:

$$\bar{x} = \frac{\sum_i \sum_j i \cdot \hat{\mathcal{H}}(i,j)}{\sum_i \sum_j \hat{\mathcal{H}}(i,j)}, \quad \bar{y} = \frac{\sum_i \sum_j j \cdot \hat{\mathcal{H}}(i,j)}{\sum_i \sum_j \hat{\mathcal{H}}(i,j)} \tag{1}$$

where the summation is over all $i \in [x_1, x_2]$ and $j \in [y_1, y_2]$.

Subsequently, we select positive point prompts for SAM from the calculated mean points $\hat{\mathcal{C}}$. For the negative point prompts, we utilize the inverse of the initial mask, specifically $1 - \hat{M}$. These prompts, automatically generated, guide SAM in generating the binary segmentation mask, $M_{\text{AP}} \in \{0,1\}^{H\times W}$.

Mask ensemble. M_{AP} and $\hat{M}$ complement each other with strengths and weaknesses. $\hat{M}$ is robust against noise like dandruff as it was trained using simulated noise. Meanwhile, M_{AP}, benefiting from SAM's superior edge detection, excels in constructing a clear boundary between hair and scalp. Therefore, to make a robust hair mask, the final binary mask, M, is derived from $\hat{M}$ and M_{AP} with the bitwise AND operation ($\boldsymbol{M} = \hat{\boldsymbol{M}} \wedge \boldsymbol{M}_{\text{AP}}$), followed by a noisy region removal post-processing step with connected-component analysis.

2.2 Scalp Condition Classification

Accurately classifying scalp disease severity from microscopic images is difficult due to the rarity of extreme cases. To address this, we introduce *DiffuseIT-M*, a diffusion-based image translation model with mask guidance that transforms a source image into various scalp conditions while preserving hair content. Building on DiffuseIT [13] and incorporating an image editing technique inspired by blended diffusion [2], *DiffuseIT-M* enables robust augmentation of underrepresented classes for improved classification.

Image translation with mask guidance. To facilitate the transfer of scalp disease characteristics while preserving hair features in our model, we utilize a comprehensive loss function, ℓ_{total} that guides the reverse process and is composed of five distinct loss components. These components consider the source image (x_{src}), the target image (x_{trg}), and the hair mask (M) as inputs. The combined loss function is defined as:

$$\ell_{total}\left(x; x_{src}, x_{trg}, M\right) = \lambda_1 \ell_{style} + \lambda_2 \ell_{content} + \lambda_3 \ell_{mask} + \lambda_4 \ell_{sem} + \lambda_5 \ell_{rng}, \tag{2}$$

where $\lambda_{i\in[1,5]}$ denotes the weights assigned to each of these loss functions.

For ℓ_{style} and $\ell_{content}$, we utilize the style and content loss functions from DiffuseIT. We employ the `[CLS]` token matching loss using DINO-ViT [3] to

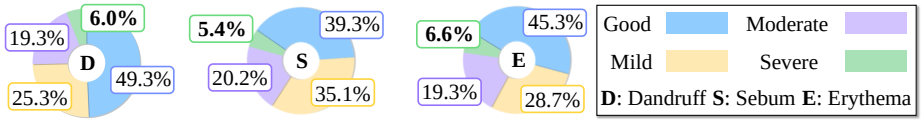

Fig. 2: Data distribution of different severity within each scalp condition.

reflect semantic information in x_{trg} and use keys of multi-head self-attention layers to preserve the content of x_{src}. Additionally, to ensure hair preservation while translating scalp styles, we construct a mask preservation loss function as:

$$\ell_{mask} = LPIPS(x_{src} \odot M, \hat{x}_0(x_t) \odot M) + ||(x_{src} - \hat{x}_0(x_t)) \odot M||_2, \quad (3)$$

where *LPIPS* denotes the learned perceptual image patch similarity metric [22] and $\hat{x}_0(x_t)$ is the estimation of the cleaned image derived from the sample x_t:

$$\hat{x}_0(x_t) = \frac{x_t}{\sqrt{\bar{\alpha}_t}} - \frac{\sqrt{1-\bar{\alpha}_t}\epsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}}. \quad (4)$$

We also include two additional losses: ℓ_{rng}, representing the squared spherical distance as proposed in [5], and ℓ_{sem}, indicating the semantic divergence loss as outlined in [13]. Using this composite loss function, ℓ_{total}, we guide the generation of the next sample step, x_{t-1}. To preserve hair details, we apply a masking approach:

$$x_{t-1} \leftarrow x_t \odot M + \left[\hat{x}_0(x_t) - \nabla x_t \ell_{total}(\hat{x}_0(x_t))\right] \odot (1 - M). \quad (5)$$

This method allows scalp style translation without extra training.

Classification strategy. Using *DiffuseIT-M*, we augment our training set by translating randomly chosen images into higher severity levels via weighted sampling, where selection probability is inversely proportional to the class's size. We fine-tune a pretrained backbone with four MLP heads, each tied to a specific loss. One head detects the presence of scalp diseases (dandruff, excess sebum, erythema), while the other three classify their severities (good, mild, moderate, severe). Our objective, ℓ_{cls}, is the sum of four losses: ℓ_{dis}(binary cross-entropy for disease presence), and ℓ_{dand}, ℓ_{seb}, ℓ_{ery}(cross-entropy for severity classification). This design enables simultaneous disease detection and severity assessment.

3 Experiments

3.1 Dataset

Despite the lack of publicly available datasets, we accessed a specialized dataset from AI-Hub [1] for classifying the severity of scalp conditions and alopecia. [†] The dataset comprises 95,910 images with a resolution of 640 × 480 pixels

[†] This dataset is provided by 'The Open AI Dataset Project (AI-Hub, S. Korea)' and is exempt from IRB approval as it does not contain any information that can identify individuals. The dataset is publicly accessible at `https://aihub.or.kr`.

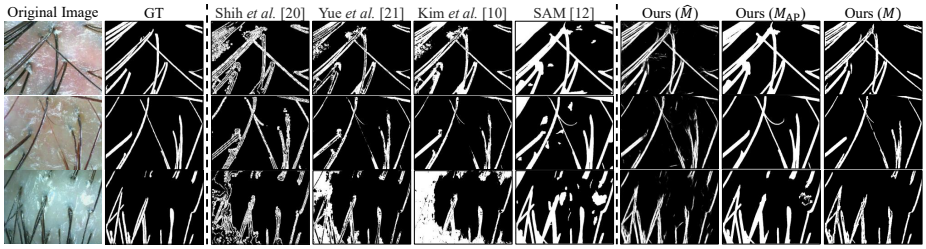

Fig. 3: Comparison of various segmentation methods on hair. "GT" represents the mask images for which we have manually annotated the pixel segmentation. Note that $\hat{M}$, M_{AP} and M are proposed in Section 2.1.

Table 1: Performance of hair segmentation on the test set.

Approach		Pixel-F1	Jaccard	Dice
Shih *et al.* [20]		0.706	0.348	0.512
Yue *et al.* [21]		0.794	0.493	0.654
Kim *et al.* [10]		0.815	0.561	0.708
SAM [12]		0.503	0.361	0.502
Ours	$\hat{M}$	0.853	0.604	0.748
	M_{AP}	0.836	0.595	0.743
	M	**0.868**	**0.649**	**0.786**

Table 2: Quantitative analysis of image-to-image translation.

Model	FID (↓)	LPIPS (↓)
DiffuseIT	138.42	0.463
AGG	141.70	0.492
Ours	**74.84**	**0.353**

from 20,000 patients. It is split into 72,342 training and 23,568 test images, with 21,703 from the training set used for validation. Dermatologists labeled each image for *dandruff*, *excess sebum*, and *erythema*, categorizing severity as *good*, *mild*, *moderate*, or *severe*. The dataset is heavily skewed toward *good* and *mild* cases (Fig. 2) and lacks segmentation labels. Thus, we manually annotated hair regions in 150 test images to evaluate our hair segmentation methods.

3.2 Hair Segmentation

Because narrow hairs are difficult to segment, many existing unsupervised methods still rely on traditional computer vision for hair-specific challenges. Accordingly, we compare prior scalp segmentation approaches [20, 21, 10] as baselines and also evaluate the foundational model SAM [12]. In addition, we conduct an ablation study on our final mask M and its intermediate versions $\hat{M}$ and M_{AP}.
Quantitative results. Table 1 reveals that our methods surpass the performance of existing hair segmentation techniques. In particular, the approach of combining the advantages of the two masks, $\hat{M}$ and M_{AP} using the bitwise AND operator in M showed the best performance. These results show the limitations of traditional computer vision techniques used in previous studies for image segmentation, revealing a lack of understanding in capturing the intricate patterns

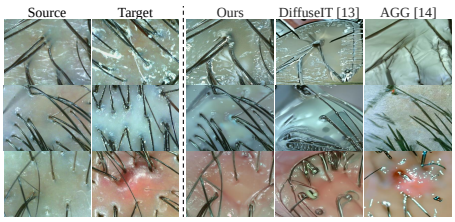

Fig. 4: Image translation results with different generative models, where the goal is preserving source hairlines while changing the scalp.

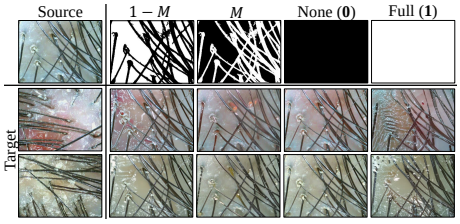

Fig. 5: Image translation results using various mask guidance. Note that our approach is guided by $1 - M$.

of hair and the scalp. Additionally, SAM was less effective for automatic segmentation when used without specific guidance.

Qualitative results. As shown in Figure 3, our approach demonstrates effective hair segmentation with robustness to noise, providing clear and accurate hair segmentation compared to previous methods. Furthermore, it shows that $\hat{M}$ faces challenges in clearly capturing hair, and it exhibits robustness against noise such as dandruff. Conversely, M_{AP} captures the hair well but is less robust to noise. Therefore, the combination of the two masks, M, demonstrates the mitigation of the drawbacks of each mask.

3.3 Synthetic Image Generation

For the evaluation of *DiffuseIT-M*, we compared our model against DiffuseIT [13] and AGG [14] as baselines for the image-to-image translation model. This evaluation demonstrates that our model not only achieves high fidelity in image-to-image translation but also effectively preserves the desired hair details.

Quantitative results. We have selected to employ the FID [8] and LPIPS [22] scores for fidelity evaluation using images from our augmentation dataset, with DiffuseIT and AGG serving as baseline models. Table 2 reveals that *DiffuseIT-M* outperforms other models in both metrics, indicating superior image fidelity. This high-quality image generation is attributed to our model's effective implementation of mask guidance.

Qualitative results. Figure 4 shows that both DiffuseIT and AGG models fail to preserve the hair content information from the source image. Furthermore, these models tended to compromise overall information and were unable to transfer the semantic information. However, our model successfully preserved hair content information and transferred the semantic information.

Effect of mask guidance. We conducted experiments to examine the impact of mask guidance on hair information preservation during image translation. As illustrated in Figure 5, our method, guided by the mask $1-M$, effectively retains hair features while successfully transferring the semantic attributes of the target image onto the scalp. In contrast, using the reverse mask, M, leads to only minor

Table 3: Performance of scalp condition classification with various augmentation methods, denoted after "+" symbol, on the test set. The second column displays the overall macro-F1 score, while the columns from the third onward show the F1 scores for each severity level of the three diseases.

Model	F1	Dandruff				Sebum				Erythema			
	macro	good	mild	moderate	severe	good	mild	moderate	severe	good	mild	moderate	severe
DenseNet	0.582	0.796	0.514	0.592	0.614	**0.554**	0.601	0.641	0.000	**0.776**	0.729	0.565	0.601
+ Gaussian Noise	0.567	0.780	0.497	0.566	0.597	0.471	0.581	0.595	0.000	0.751	0.712	0.614	0.635
+ AugMix	0.525	0.789	0.501	0.589	0.000	0.504	0.588	0.634	0.000	0.743	0.718	0.596	0.585
+ DiffuseIT	0.608	0.809	0.482	0.604	0.650	0.536	0.613	0.625	0.202	0.774	0.740	0.621	0.639
+ AGG	0.610	0.811	0.480	0.591	0.654	0.518	0.598	0.612	0.300	0.771	**0.740**	**0.629**	0.621
+ Ours	**0.636**	**0.820**	**0.541**	**0.625**	**0.665**	0.536	**0.617**	**0.641**	**0.430**	0.758	0.734	0.621	**0.641**
EfficientFormerV2	0.569	0.795	0.417	0.598	0.628	0.526	0.565	0.628	0.000	0.772	0.709	0.623	0.569
+ Gaussian Noise	0.562	0.780	0.477	0.566	0.633	0.460	0.585	0.550	0.000	0.742	0.714	0.598	0.637
+ AugMix	0.577	0.789	0.494	0.592	0.635	0.519	0.593	0.623	0.000	0.746	0.724	0.620	0.590
+ DiffuseIT	0.596	0.798	0.441	0.598	0.632	0.526	0.595	0.606	0.236	0.766	0.715	0.612	0.621
+ AGG	0.610	0.801	0.509	0.604	0.626	0.511	0.583	0.608	0.300	**0.787**	0.736	0.624	0.628
+ Ours	**0.635**	**0.807**	**0.529**	**0.619**	**0.669**	**0.535**	**0.613**	**0.632**	**0.406**	0.781	**0.738**	**0.639**	**0.648**

alterations in scalp color from the target image, with a notable transfer of hair semantic information from the target. When no mask ($\mathbf{0}$) is applied, the translation results in minimal color change, failing to transfer conditions like dandruff from the target. Conversely, with a full mask ($\mathbf{1}$), both hair and scalp features are subjected to changes. This differentiation in results highlights the importance of mask guidance in preserving specific image features, demonstrating the versatility of our approach in handling different translation objectives.

3.4 Scalp Condition Classification

To demonstrate the effectiveness of our augmentation method using generated images, we employed two different models as the classification backbone: DenseNet [9] as a CNN and EfficientFormerV2 [15] as a Transformer. As summarized in Table 3, we compared only those augmentation methods that preserve a one-to-one correspondence between each image and its original, unambiguous set of condition labels— Gaussian noise, AugMix [7], DiffuseIT [13], and AGG [14]. Our approach, which specifically employs *DiffuseIT-M*, achieved the highest performance in both models. Notably, classifying the *severe* sebum class proved to be especially challenging when using non-generative augmentation methods. This difficulty arises primarily due to the extreme scarcity of samples for this class. The augmentation of the training dataset with generative models led to enhanced performance compared to the baseline. Our model, in particular, exhibited superior accuracy compared to DiffuseIT and AGG, which struggled to preserve the essential information of the hair effectively. This underscores the significance of incorporating both the scalp style details and the hair content information in the scalp disorder classification.

4 Conclusion and Discussion

In this work, we introduced ScalpVision, a diagnostic system designed for a complete evaluation of scalp health. Our approach combines label-free hair segmentation – based on a naive segmentation model and a foundation segmentation model (SAM) – with diffusion-based data augmentation to address data imbalance and preserve critical hair features. However, scalp disorders are influenced by both scalp conditions and hair information. Thus, we plan to utilize hair information to extend our research to conditions like alopecia beyond the three scalp diseases. We envision ScalpVision as a meaningful step towards a generalized diagnostic system for dermatological applications.

Acknowledgments. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00354218 and RS-2024-00353125). Junhyug Noh was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2022-00155966).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Bibliography

[1] AI Hub: Scalp and Hair Follicle Image Dataset (2020), https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=207, [Online; accessed 25 February 2025]

[2] Avrahami, O., Lischinski, D., Fried, O.: Blended diffusion for text-driven editing of natural images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18208–18218 (2022)

[3] Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., Joulin, A.: Emerging properties in self-supervised vision transformers. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 9650–9660 (2021)

[4] Chang, W.J., Chen, L.B., Chen, M.C., Chiu, Y.C., Lin, J.Y.: Scalpeye: A deep learning-based scalp hair inspection and diagnosis system for scalp health. IEEE Access **8**, 134826–134837 (2020)

[5] Crowson, K., Biderman, S., Kornis, D., Stander, D., Hallahan, E., Castricato, L., Raff, E.: Vqgan-clip: Open domain image generation and editing with natural language guidance. In: European Conference on Computer Vision. pp. 88–105. Springer (2022)

[6] Elewski, B.E.: Clinical diagnosis of common scalp disorders. In: Journal of Investigative Dermatology Symposium Proceedings. vol. 10, pp. 190–193. Elsevier (2005)

[7] Hendrycks, D., Mu, N., Cubuk, E.D., Zoph, B., Gilmer, J., Lakshminarayanan, B.: Augmix: A simple data processing method to improve robustness and uncertainty. arXiv preprint arXiv:1912.02781 (2019)

[8] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: Gans trained by a two time-scale update rule converge to a local nash equilibrium. Advances in neural information processing systems **30** (2017)

[9] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)

[10] Kim, H., Kim, W., Rew, J., Rho, S., Park, J., Hwang, E.: Evaluation of hair and scalp condition based on microscopy image analysis. In: 2017 International conference on platform technology and service (PlatCon). pp. 1–4. IEEE (2017)

[11] Kim, J.H., Kwon, S., Fu, J., Park, J.H.: Hair follicle classification and hair loss severity estimation using mask r-cnn. Journal of Imaging **8**(10), 283 (2022)

[12] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. arXiv preprint arXiv:2304.02643 (2023)

[13] Kwon, G., Ye, J.C.: Diffusion-based image translation using disentangled style and content representation. In: The Eleventh International Conference on Learning Representations (2022)

[14] Kwon, G., Ye, J.C.: Improving diffusion-based image translation using asymmetric gradient guidance. arXiv preprint arXiv:2306.04396 (2023)
[15] Li, Y., Hu, J., Wen, Y., Evangelidis, G., Salahi, K., Wang, Y., Tulyakov, S., Ren, J.: Rethinking vision transformers for mobilenet size and speed. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 16889–16900 (2023)
[16] Panjwani, S.: Early diagnosis and treatment of discoid lupus erythematosus. The Journal of the American Board of Family Medicine **22**(2), 206–213 (2009)
[17] Pratt, C.H., King, L.E., Messenger, A.G., Christiano, A.M., Sundberg, J.P.: Alopecia areata. Nature reviews Disease primers **3**(1), 1–17 (2017)
[18] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O.R., Jagersand, M.: U2-net: Going deeper with nested u-structure for salient object detection. Pattern recognition **106**, 107404 (2020)
[19] Seo, S., Park, J.: Trichoscopy of alopecia areata: hair loss feature extraction and computation using grid line selection and eigenvalue. Computational and Mathematical Methods in Medicine **2020** (2020)
[20] Shih, H.C.: An unsupervised hair segmentation and counting system in microscopy images. IEEE Sensors Journal **15**(6), 3565–3572 (2014)
[21] Yue, G., Ji, C., Yang, Y., et al.: Hair counting method based on image processing technology. Journal of Artificial Intelligence Practice **4**(1), 23–29 (2021)
[22] Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)
[23] Zhang, T.Y., Suen, C.Y.: A fast parallel algorithm for thinning digital patterns. Commun. ACM **27**(3), 236–239 (mar 1984). https://doi.org/10.1145/357994.358023, https://doi.org/10.1145/357994.358023