Weakly-Supervised Hair SEM Microscope Image Segmentation Using a Priori Structure Information

Xiaohu Liu, Thérèse Baldeweck, Pierre Dupuis, Thomas Bornschloegl, Etienne Decencière, Beatriz Marcotegui

Image Anal Stereol 2024;43:259-275 doi: 10.5566/ias.3406
Original Research Paper

WEAKLY-SUPERVISED HAIR SEM MICROSCOPE IMAGE SEGMENTATION USING A PRIORI STRUCTURE INFORMATION

XIAOHU LIU[1], THÉRÈSE BALDEWECK[2], PIERRE DUPUIS[2], THOMAS BORNSCHLOEGL[2], ETIENNE DECENCIÈRE[1] AND BEATRIZ MARCOTEGUI✉,[1]

[1]Mines Paris, PSL Research University, center for Mathematical Morphology, 35 rue St Honoré, 77300 Fontainebleau, France. e-mail: name.surname@minesparis.psl.eu , [2]L'Oréal Research and Innovation, 1 avenue Eugène Schueller, 93600 Aulnay-sous-Bois, France. e-mail: name.surname@loreal.com
(Received October 1, 2024; accepted October 28, 2024)

ABSTRACT

The analysis of microscopic images of hair has a wide range of applications in the domains of cosmetics, healthcare and forensics. The segmentation of the hair represents the initial step of automatic large-scale quantitative analysis of microscopic images of hair. It ensures that subsequent quantitative measurements are performed in an appropriate area corresponding to the hair, avoiding artifacts near the boundaries. This process can be time-consuming, tedious and susceptible to subjective errors when conducted by a human operator.
Deep learning methods represent a promising solution; however, obtaining pixel-level accurate masks is a costly process. This paper presents a novel weakly supervised pipeline for the segmentation of hair SEM (Scanning Electron Microscope) microscopic images, which requires only simple image-level annotations for training. The proposed method incorporates the Radon transform, the Sobel operator and a novel Boundary Discrimination Module (BD-module) for the estimation of the presence of boundaries.
The proposed pipeline was evaluated on a recently collected hair SEM dataset (429 images). Furthermore, it is benchmarked with methods including Unet and Segment Anything Model (SAM). The results demonstrated a mean Hausdorff Distance improvement of over 30% and a standard deviation improvement of over 50% in comparison with the Unet and SAM. Moreover, we proposed additional refinement modules to address boundary nonlinear cases and conducted Grad-CAM analysis to enhance the interpretability of the BD-module. Additionally, we proposed a novel quality estimation metric based on gradient map for self-quality assessment. The SEM hair dataset is accessible to the research community in an open-source format.
Keywords: Hair microscopic segmentation, Radon transform, SEM segmentation, weakly-supervised segmentation.

INTRODUCTION

One of the preliminary steps in the automatic analysis of hair micrographs is the segmentation of the Region of Interest (ROI) (Chu *et al.*, 2020). This serves as the foundation for subsequent studies on morphological texture analysis and classification of damage types, and is crucial for the accurate assessment of hair properties (Chu *et al.*, 2020; Shih, 2014; Lee *et al.*, 2017; Kim *et al.*, 2010; Tomes *et al.*, 2007; Lee *et al.*, 2016). For example, due to the 3D nature of hair, the same quantitative measures, such as scale size, may be subject to significant variation in close proximity to the boundary areas (Chu *et al.*, 2020; Tomes *et al.*, 2007). Therefore, segmentation represents an essential preliminary step in this process. The objective is to remove the background portion that contains noisy signals and retain only the foreground portion. For humans, the distinction between the foreground and background in hair micrographs is readily apparent under ideal Scanning Electron Microscope (SEM) acquisition conditions. However, in the presence of low contrast or boundary artefacts, the accurate distinction may prove challenging. The foreground is delineated by approximately parallel hair boundaries and contains textured parts such as the cuticle and cortex, which are usually brighter relative to the background. The background is primarily comprised of the carrier plate and the adhesive utilized to affix the hair in place. On occasion, the adhesive may act as artefacts that present a challenge for the delineation of boundaries.

Though user-defined ROI could be used for small-scale image analysis (Tomes *et al.*, 2007; Kim *et al.*, 2010; Chu *et al.*, 2020; Lee *et al.*, 2016), when confronted with large-scale datasets (Zhang *et al.*, 2021; Man *et al.*, 2021), blurred and dark boundaries resulting from disparate acquisition conditions, and irregular boundaries due to cuticle buckling, manual annotation is inefficient, costly, and prone to subjective errors. Consequently, it is imperative to identify automated foreground segmentation methods that are both efficient and accurate.

Deep learning based methods have been extensively studied in recent years, and could provide promising solution for such a problem. However, one of the challenges commonly encountered in the field

of deep learning is the necessity for a considerable number of pixel-level ground-truth segmentation labels for effective training. As the process of accurate annotation can be time consuming, weakly-supervised segmentation employs a reduced number of cost-effective image-level, bounding box, click, or scribble labels for the generation of pseudo-masks, which could then be refined to obtain the final masks (Ahn and Kwak, 2018; Huang *et al.*, 2018; Li *et al.*, 2022; Kulharia *et al.*, 2020; Chen *et al.*, 2021; Lee and Jeong, 2020). Besides, the recent advent of pre-trained foundation models for segmentation tasks, exemplified by the Segment Anything Model (SAM) (Kirillov *et al.*, 2023), have offered a promising solution to the problem of zero-shot automatic hair microscopy image foreground segmentation. SAM has demonstrated excellent performance in our hair microscopic images obtained under ideal acquisition conditions when a single center point is used as the foreground prompt. However, the segmentation performance of the single centroid as the prompt is sensitive to texture changes (e.g., cuticle partially detached and cortex partially exposed). Furthermore, the mask generated under the single centroid condition exhibits ambiguity that may potentially be mitigated by incorporating a larger number of prompt points in order to improve its overall robustness. This issue will be addressed in the subsequent sections.

This paper presents a novel weakly-supervised pipeline for the segmentation of hair micrographs, which necessitates only image-level boundary existence annotations for training. The method employs the Sobel operator, the Radon transform, and a Convolutional Neural Network (CNN) based Boundary Discrimination Module (BD-module) for pseudo-mask generation. Furthermore, we show how to use Unet (Ronneberger *et al.*, 2015) and SAM for boundary refinement and extension to encompass curved boundary cases. The BD-module is incorporated as the hair micrographs could comprise a highly variable range of boundary cases. As illustrated in Fig.1, the SEM images include cases of two edges (a), one edge (b), hair occupying the whole image (c), low light (d), multiple hairs (e) or artefacts at boundaries (f).

The principal contributions of our work can be summarized as follows:

- We propose a novel weakly-supervised hair microscopic image segmentation pipeline that combines classical image processing techniques with a deep learning model, requiring only image-level annotations. The proposed pipeline was evaluated on a SEM hair dataset, with benchmarks set with Unet and SAM. The results demonstrated that the proposed method yielded superior outcomes in terms of performance and interpretability compared with benchmarks.
- A novel segmentation quality estimation metric based on detected gradient map is put forth as a means of automatically identifying erroneous results.
- The microscopic SEM image dataset and the corresponding ground truth foreground masks are made available in an open-source format to allow further research in the field.

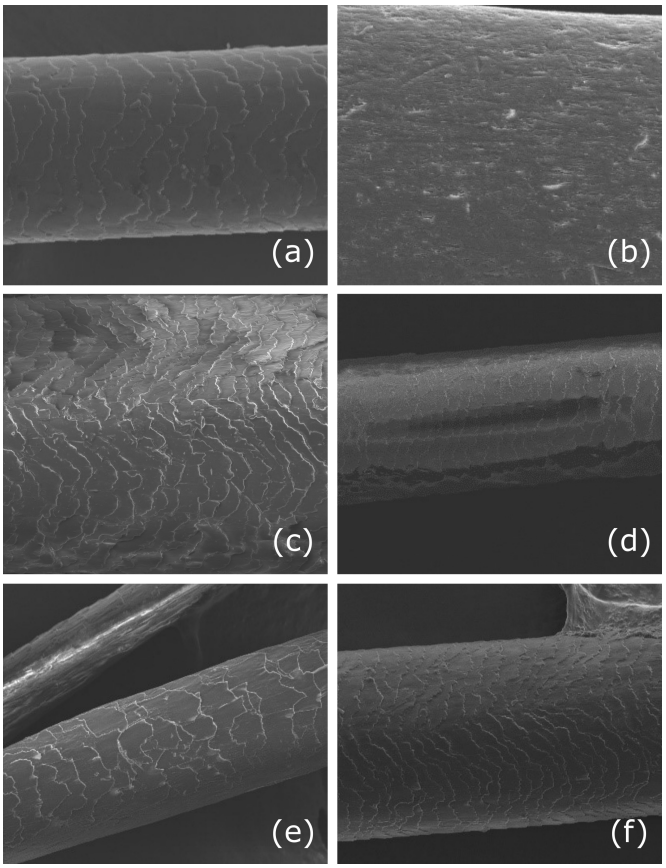

Fig. 1. *Examples of SEM microscopy images from the dataset, including cases with two clear borders (a), with a single (incomplete) border (b), without borders (c), with a border of low contrast compared to the background (d), with multiple hairs (e), with a border containing artefacts such as glue (f).*

RELATED WORK

HAIR MICROSCOPIC IMAGE ANALYSIS

The analysis of microscopic images of hair represents a valuable non-invasive analytical method for the rapid estimation of hair and scalp health and damage levels, as well as changes in hair surface

properties caused by external influences such as air pollution (Coroaba *et al.*, 2020; Tomes *et al.*, 2007; Kim *et al.*, 2010; Lee *et al.*, 2016; Chu *et al.*, 2020; Zhang *et al.*, 2021; Man *et al.*, 2021; Shih, 2014; Galliano *et al.*, 2017; 2023). Coroaba *et al.* (2020) performed a variety of physicochemical treatments on hair to obtain images for morphological analysis to study the effects of baldness on the morphology and nature of hair. To obtain information such as the height of the cuticle scales, Tomes *et al.* (2007) analyzed a reconstructed three-dimensional model of hair fibers obtained from multiple 2D SEM images. In order to establish a standard and objective scoring system for the quantitative assessment of hair damage, Kim *et al.* (2010) first proposed a classification system with 5 damage scores on SEM and TEM images, which was then further developed into a 12-point scale system by Lee *et al.* (2016). Whereas previous scoring methods are based on the presence/absence of certain morphological features (i.e. lifting of cuticle edges, exposure of cortex, etc.), Chu *et al.* (2020) proposed a set of quantitative measures to assess hair damage without the use of specific features, including pixel brightness asymmetry histograms and surface roughness. Although the proposed protocol aims to automate the damage assessment process, it is based on user-defined ROIs, which could act as a bottleneck for the automatic analysis of large scale datasets. The effects of particulate matter from air pollution on hair surfaces have also been investigated using SEM images (Galliano *et al.*, 2017; 2023). Nevertheless, the use of user-defined ROIs in the studies may still present a challenge for subsequent research.

The accelerated advancement and extensive use of deep learning in the domain of computer vision in recent times have yielded novel conceptual frameworks for the automated analysis of hair microscopy images (Zhang *et al.*, 2021; Man *et al.*, 2021). To automatically quantify the degree and area of hair damage, Zhang *et al.* (2021) proposed HDM-NET based on MobileNet and SVM to classify SEM hair micrographs into three categories: low, moderate and high damage. In the same line of work, Man *et al.* (2021) further collected a larger SEM microscopy image dataset containing in total 15,000 images, and proposed SACN-Net combining convolutional neural network with an attention mechanism module, which achieved 98% accuracy on the three categories. Nevertheless, the classification into weakly, moderately and highly damaged is coarse-grain, which limits the ability to perform a more precise classification of damage types (e.g., cortex appearance ratio or cuticle lift degree).

Hair segmentation serves as an important step for subsequent analysis of the hair (e.g. counting, estimating hair width, etc.) (Shih, 2014; Lee *et al.*, 2017). More precisely, in the context of hair micrograph segmentation and counting, Shih (2014) proposed the use of multiscale Hough transform and parallel line bundling algorithms to address the issue of hair length variations, as well as the presence of surface artefacts and the overlapping and occlusion of multiple hairs. Still in the background of hair segmentation, to solve the problem of loss of contrast between hair and background colors and artefact of background (skin) texture in micrographs, Lee *et al.* (2017) proposed a method to generate hair mask by calculating Edge Density and Mean Branch Length with adaptive thresholds, and further refining the mask by applying K-NN classification. The aforementioned methods address the same hair micrograph segmentation task but differ in their focus.

While the existing approaches concentrate on hair segmentation and artefact processing at low resolution (μm/pixel), to the best of our knowledge, our work is the first to explore the intricacies of hair SEM micrograph segmentation at higher resolution ($\leq$100 nm/pixel). For comparison, the occlusion and overlapping of hair are less prevalent in our task. However, the surface texture of the hair is more visible. Additionally, our task is confronted with the issues of blurred edges and low contrast resulting from alterations in the light conditions during acquisition, as well as the impact of artefacts such as glue used to fix the hair on the supporting plate. An example of this artifact is illustrated in the example (f) in Fig. 1.

WEAKLY-SUPERVISED SEGMENTATION

The creation of pixel-wise accurate mask labels is a costly and error-prone process. The search for cheaper and more efficient types of annotation has therefore been widely investigated. Weakly-supervised segmentation aims to reduce the dependence on pixel-level labels. Existing methods include image-level class labels based methods, bounding box labels based methods, click (point) or scribbles annotations based methods (Ahn and Kwak, 2018; Huang *et al.*, 2018; Li *et al.*, 2022; Kulharia *et al.*, 2020; Chen *et al.*, 2021; Lee and Jeong, 2020). Methods using image-level class label-based methods typically generate Class Activation Maps (CAMs) (Selvaraju *et al.*, 2017) through supervised training of a classification network. Subsequently, the CAMs are refined to obtain pseudo-masks which will be further used to train the segmentation networks. Although class labels are much less costly compared to pixel-level, the method is an ill-posed problem since the class labels themselves do not contain spatial location information and it is difficult to obtain accurate pseudo-masks (Li *et al.*, 2022).

In the field of microscopic image analysis, image-level labels, point annotations and scribble labels have been recently explored for cell, mitochondria, nucleus, histopathological and brain neuron segmentations (Bilodeau *et al.*, 2022; Kniesel *et al.*, 2023; Li *et al.*, 2023; Qiu *et al.*, 2024; Zhao and Yin, 2020; Obikane and Aoki, 2020; Dong *et al.*, 2020; Lee and Jeong, 2020). More precisely, Bilodeau *et al.* (2022) proposed MICRA-Net, which addresses auxiliary tasks including semantic segmentation, detection, and enumeration through image-level annotations and CAMs. Qiu *et al.* (2024) explored further reduction of the point label number by proposed WDA-Net, which needs only sparse points annotations for a subset of the whole image. Li *et al.* (2023) proposed a novel approach that integrates generative models and point labels. The proposed method demonstrated comparable performance to the baseline method, which was trained with real images. Kniesel *et al.* (2023) used domain specific image-level labels (absence/presence of virus) to train a binary classification network and used CAMs maxima in order to initialize the detection center, which greatly reduced the cost of labelling.

The approach presented in this paper is aligned with image-level labels based methods. It uses image-level labels to produce a pseudo-mask and a refinement network for further refinement. However, our proposed pipeline does not use CAMs. It incorporates instead the trained CNN into a classical image processing pipeline, thereby employing domain-specific knowledge and circumventing the inaccurate pseudo-mask produced by CAMs.

DATASET

The dataset used in this study is a hair SEM image dataset with corresponding ground truth foreground mask[1] annotated with VIA annotator (Dutta and Zisserman, 2019) and a third part algorithm. The dataset comprises a total of 429 micrographs[2], divided into two sub-datasets, denoted as Base1 and Base2.

The Base1 dataset comprises 240 microscopic images acquired with a HITACHI S3400N SEM microscope with a resolution of 1280x960 pixels (99 nm/pixel). The Base2 dataset comprises 189 images acquired with an FEI QUANTA 400F SEM microscope with a resolution of 2048 x 1887 pixels (62 nm/pixel).

Both datasets encompass a variety of damage types, hair foreground percentages, acquisition conditions, and background artefacts, which are illustrated in Fig. 1.

APPROACH

This section presents a novel pipeline for foreground segmentation of hair microscopy images. The input microscopic image I is primarily convolved by the vertical Sobel operator, denoted as G_y, to calculate the horizontal boundary image S_h^I:

$$S_h^I = G_y * I \tag{1}$$

The rationale for using horizontal boundary S_h^I as input is that all SEM images in the dataset exhibit horizontal hair positions, and that this approach yields superior performance compared to experiments conducted with the original image I as input.

The pipeline applied on the S_h^I comprises three modules, as illustrated in Fig. 2.

- **Step 1: Boundary Discrimination (BD) module**: One of the challenges associated with this dataset is the occurrence of instances where the hair is particularly thick, resulting in its overflow beyond the confines of the image, or alternatively, instances where the border is only partially visible. The annotation process at the image level is relatively straightforward. It involves the application of binary labels to indicate the presence or absence of borders along the upper and lower edges. Subsequently, a CNN is trained to predict this information.
- **Step 2: Radon transform-based pseudo-mask module**: The lines corresponding to hair boundaries (orientation and distance to center) are determined through the Radon transform. A pseudo-mask $\hat{M}$ is generated by combining the aforementioned predicted boundary existence variables.
- **Step 3: Refinement module**: The pseudo-mask $\hat{M}$ is employed for the purpose of further training a deep learning model, such as Unet, in a supervised manner or for randomly sampling prompt points for the pre-trained Segment Anything model (SAM). This process serves to further refine the mask boundary to encompass cases where the boundary is not straight.

[1]The ground truth masks are composed of two parts: 84 images are annotated with VIA annotator and the rest are from a third part algorithm. All the ground truth masks are validated by experts in L'Oréal research center

[2]The dataset is available at: https://people.cmm.minesparis.psl.eu/files/databases/Loreal_SEM_Hair_2024/

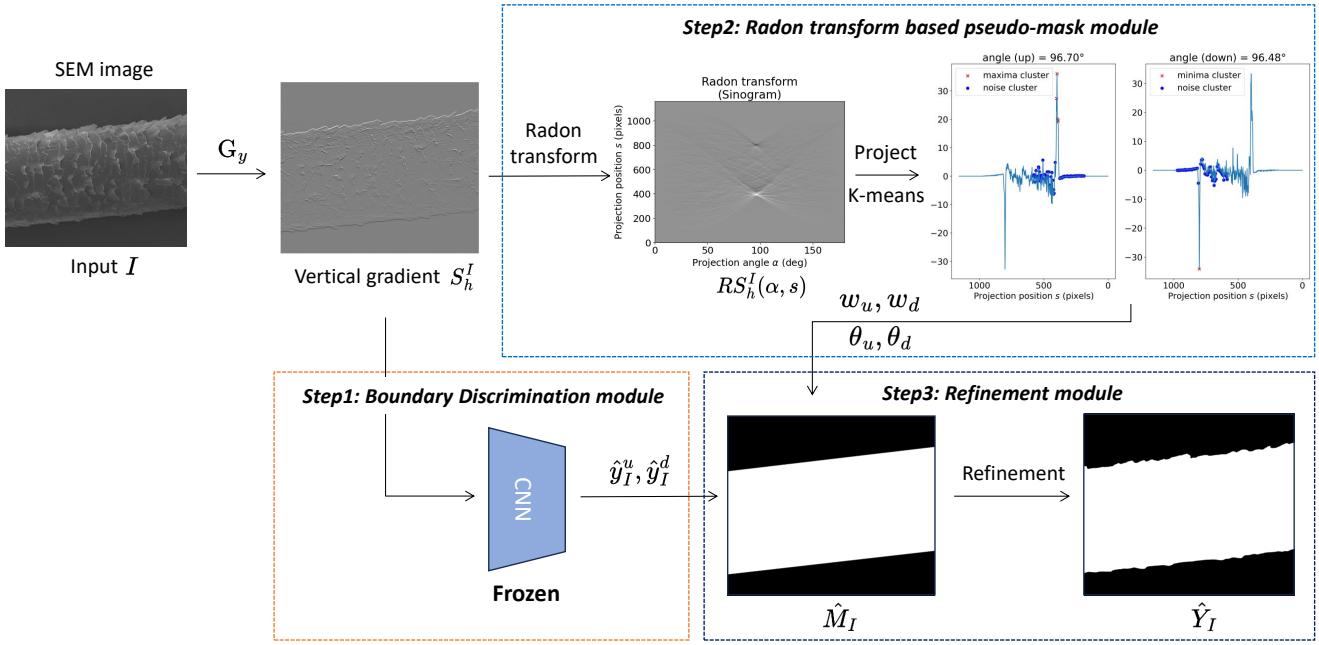

Fig. 2. *Our proposed pipeline for hair SEM image foreground segmentation. The Boundary Discrimination Module employs a CNN to estimate the existence of upper and lower boundaries (Step1). The Radon transform based pseudo-mask module (Step2) generates straight line edge pseudo-mask based on the Radon transform and the Sobel operator. The Refinement module (Step3) further refines the boundary details and extends the method to non-linear boundary cases by employing SAM or Unet.*

The following section provides a detailed description of the three modules.

STEP 1: BOUNDARY DISCRIMINATION (BD) MODULE

One of the key aspects of foreground segmentation in hair micrographs is the identification of upper and lower boundaries. This is evident when the image acquisition conditions are optimal, allowing for the clear delineation of boundaries through simple image/signal processing. However, when the conditions of image acquisition change (e.g., low light, multiple hair segments at the same time, the presence of impurities such as glue at the boundaries, or incomplete hair boundaries of short lengths), the simple approach fails. Recently, Liu *et al.* (2024) demonstrated the effectiveness of incorporating CNN in a traditional morphological pipeline for adaptive parameters estimation. Inspired by this work, we propose to integrate a CNN into a classical image processing pipeline for boundary existence discrimination to handle difficult cases where hair boundaries are barely visible or have artefacts. The details of the architecture of this module are detailed in Section 5.1.

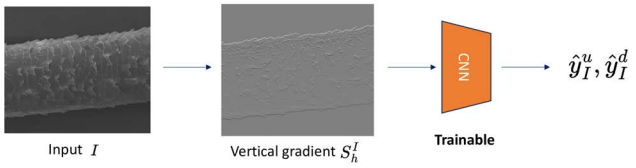

Fig. 3. *Training stage of Boundary Discrimination network.*

Fig. 3 illustrates the flow of the BD-module training. the vertical gradient image S_h^I is processed by the CNN to generate variables $\hat{y}_I^u$ and $\hat{y}_I^d$ representing the estimated probabilities for the existence of the upper and lower boundaries, respectively, after the application of the sigmoid activation function:

$$\hat{y}_I^{u,d} = Sigmoid(CNN_\theta(S_h^I)) \tag{2}$$

Subsequently, the Binary Cross Entropy (BCE) loss function is calculated with the manually annotated ground truth labels denoted as y_I^u and y_I^d.

In the inference stage, for new microscopic images that are not included in the training set, the pre-trained network with fixed network parameters computes the upper and lower boundary existence probabilities $\hat{y}_I^u$

and $\hat{y}_I^d$ through S_h^I. The boundary is considered present if probability is more than 0.5.

STEP 2: RADON TRANSFORM BASED PSEUDO-MASK MODULE

Investigated in macro-scale microscopy hair segmentation (Shih, 2014), the Hough transform and its counterpart, the Radon transform, are closely related (or equivalent) (van Ginkel *et al.*, 2004). However, the Radon transform has been shown to be more robust to noise than its Hough transform counterpart (Bhaskar and Werghi, 2011) and has the advantage of being more intuitive with a more solid mathematical basis (van Ginkel *et al.*, 2004). Thus we use the Radon transform for better robustness when dealing with background artefacts such as glue.

Given that the boundaries of hair micrographs are predominantly straight (or nearly straight), integrating the vertical gradient (assuming a horizontal hair pose) in the same direction as the boundaries results in the emergence of significant maxima and minima, which correspond to the two boundary positions (if they are present). From the sinogram's extreme points, the direction of the boundary and the distance from the center point can be determined, allowing the construction of a linear boundary pseudo-mask. The procedure is depicted in Fig. 2.

The Sinogram[3] RS_h^I for image S_h^I is given by:

$$\begin{aligned} RS_h^I(\alpha, s) &= \int_{-\infty}^{\infty} S_h^I(x(z), y(z))\, dz \\ &= \int_{-\infty}^{\infty} S_h^I\big(z\sin\alpha + s\cos\alpha, \\ &\qquad (-z\cos\alpha + s\sin\alpha)\big)\, dz \end{aligned} \tag{3}$$

where α is the projection angle and s the projection pixel. As illustrated in Fig. 2, when the projection angle is aligned with a hair boundary, an extremum is obtained in the sinogram.Thus, the upper boundary can be estimated by determining the position where RS_h^I achieves its maximum value:

$$\alpha_{max}, s_{max} = {}_{(\alpha,s)\in[0,\mathscr{A}]\times\left[\frac{\mathscr{S}}{2},\mathscr{S}\right]} RS_h^I \tag{4}$$

where $\mathscr{A}$ and $\mathscr{S}$ denote the maximum range for angle and projection pixel axis for RS_h^I respectively. The choice of the search range for s in $[\frac{\mathscr{S}}{2}, \mathscr{S}]$ is linked to the acquisition condition: the hair must contain the center of the image. Therefore, the upper boundary is located within the lower half of RS_h^I.

However, the projection pixels s_{max} are inadequate for determining the distance of the boundary from the center point. This is due to the potential influence of artefacts, such as long physical damage gaps or optical texture artefacts resulting from acquisition conditions in the middle of the foreground with the same orientation as the boundary. These factors could easily compromise the precision of the measurement. Therefore, we propose the use of the K-means algorithm to separate local maxima (minima) clusters, which contain candidate maxima with boundary distances, and local noise clusters, which represent the texture artefacts:

$$\hat{\mathscr{M}}_I^u = {}_{s\in\left[\frac{\mathscr{S}}{2},\mathscr{S}\right]} \left(\mathscr{G}_{\sigma=2}(RS_h^I(\alpha_{max}, s))\right) \tag{5}$$

$$\mathscr{M}_I^u, \mathscr{N}_I^u \leftarrow (\hat{\mathscr{M}}_I^u) \tag{6}$$

where $\mathscr{G}_{\sigma=2}$ denotes Gaussian filter with σ set to 2. The final point will be selected as the one in the cluster of maxima and minima points which is furthest from the center point. The purpose of this selection is to reduce the effect of texture and light artifacts:

$$w_u = {}_{s\in\mathscr{M}_I^u} |s - \frac{\mathscr{S}}{2}|, \quad \theta_u = \alpha_{max} \tag{7}$$

The angle and distance from center for lower boundary θ_l, w_l are determined in a similar way, but the search interval is constrained by α_{max}:

$$\alpha_{min}, s_{min} = {}_{(\alpha,s)\in[\alpha_{max}-\delta,\alpha_{max}+\delta]\times\left[0,\frac{\mathscr{S}}{2}\right]} RS_h^I(\alpha, s) \tag{8}$$

The value of δ is set to 7°, which represents the degree of tolerance for parallelism observed experimentally.

Ultimately, by combining the estimated boundary existence variables $\hat{y}_I^u$ and $\hat{y}_I^d$ with the angles/distances for the upper and lower boundaries $\theta_u, \theta_l, w_u, w_l$, we can calculate the boundaries' linear functions, which determine the regions of the foreground mask.

STEP 3: REFINEMENT MODULE

The pseudo-mask obtained from the Radon transform is capable of accurately representing a straight boundary. However, this method is limited in its ability to handle cases where the hair radius changes or the boundaries are irregular. Consequently, further refinement modules are proposed based on the Unet and pre-trained SAM, with the objective of further refining the boundary details of the aforementioned method and extending it to the case where the boundary is not straight.

[3]The angular resolution is set to be proportional to the SEM width, with a rescale factor of 0.4 for better calculation efficiency. This results in an angular resolution of approximately 0.21°/pixel for Base1 and 0.35°/pixel for Base2.

1) Unet-based refinement module

Following Bilodeau *et al.*; Kniesel *et al.*; Ahn and Kwak; Huang *et al.*; Li *et al.*, we take the aforementioned pseudo-mask as labels to train a Unet in a supervised manner, and use the inference output as refinement. We applied a fill holes operator to the resulting mask and only keep the largest connected component.

2) SAM-based refinement module

As a foundational model for promptable segmentation, SAM is capable of producing high-quality masks by employing points, box regions, and text as prompts. The pseudo-mask is employed as a sampling region for the foreground and background point prompt sampling.

Foreground prompt points sampling. The initial step is to perform an erosion transform on the pseudo-mask with a square structuring element[4]. This is followed by a random sampling of the eroded pseudo-mask to obtain the positive foreground points. The rationale behind the erosion transform is to ensure that the sampled intervals are included in the foreground.

Background prompt points sampling. In a symmetrical manner, a dilation transformation is applied to the pseudo-mask, and the dilated pseudo-mask is inverted and randomly sampled to obtain the background points.

Mask refinement and post-processing. Finally, the foreground and background sampling points mentioned above are fed into the pre-trained SAM as prompts, as well as the original microscopic images, to generate the refined masks. Same as Unet, a fill holes operation is applied to the resulting mask and finally only its largest connected component is kept.

QUALITY ESTIMATION METRIC

While our proposed methods show promising performance at the dataset level, it is also interesting to provide confidence scores at the instance level, allowing for easy identification and prioritization of cases with a higher propensity for error.

In this section, we propose a new metric for automatic evaluation and uncertainty estimation based on the detected hair boundary gradient map.

Upon disposal of ground truth masks, the evaluation metrics like detection percentage (DP) and quality percentage (QP) can be calculated for the purpose of comparing disparate segmentation results, as exemplified by Angulo *et al.* (2009) for hyperspectral image segmentation. However, these methods are constrained when applied to large-scale datasets where ground truth masks are not available. To this end, unsupervised quality assessment methods have been investigated in the field of image segmentation as a means of addressing this issue (Wang *et al.*, 2020). Especially, in recent times, the estimation of uncertainty scores based on predicted network entropy has been the subject of investigation with a view to unsupervised automatic boundary segmentation quality estimation (Martins da Cruz *et al.*, 2024). The method is based on the deep network outputs, which limits the range of applications. Therefore, we propose an alternative metric to directly evaluate the quality of the detected boundary gradient. A high contrast boundary, as well as a clear and clean boundary, should have a high mean and a low variance of gradient. Conversely, a low contrast boundary or a boundary with artefacts could result in a discontinuity of the gradient, leading to less homogeneous profiles.

Given that hair boundaries are typically not perfectly straight and may exhibit cuticle warping, we compute the mask boundaries $\overline{Y}$ of width N pixels by calculating morphological gradient (Rivest *et al.*, 1992) of the prediction masks $\hat{Y}$. This is achieved by subtracting the erosion result from the dilation result of $\hat{Y}$. In order to eliminate outlier pixels with extreme values and enhance robustness, we apply area opening (Vincent, 1993) to both the original image and the vertical gradient map. The filtered image $\tilde{I}$ can be obtained with:

$$\tilde{I} = \gamma_\lambda^a(I) \tag{9}$$

With γ_λ^a denotes area opening operation with area threshold λ. The area threshold λ is set as 128×128 pixels. The area size is chosen based on image resolution and outlier regions. We then obtain the estimated boundary gradient $\overline{S_h^{\tilde{I}}}$ (Local contrast in Fig.5) by computing the dot product with the dilated vertical gradient image $S_h^{\tilde{I}}$.

$$\begin{aligned} \overline{Y} &= \hat{Y} \oplus B - \hat{Y} \ominus B \\ \overline{S_h^{\tilde{I}}} &= \overline{Y} \cdot \gamma_\lambda^a(|S_h^{\tilde{I}}| \oplus B) \end{aligned} \tag{10}$$

where $\oplus$ and $\ominus$ denote morphological dilation and erosion respectively, B denotes the structuring element with size W, thus $N = 2W$. In order to circumvent the potential for irregular cuticles at boundaries to yield

[4]In the experimental sections, the size of the structuring element was tested at 40 pixels. The number of foreground and background sampling points was set to 40 and 20, respectively.

extreme values, a gradient profile along the x axis $\overline{\mathrm{VP}_u}$ is obtained by averaging the gradient for each column according to the following equation:

$$\overline{\mathrm{VP}_u}(x) = \frac{\sum_{i=0}^{\mathscr{H}/2} \overline{S_h^I}(i,x)}{\sum_{i=0}^{\mathscr{H}/2} \overline{Y}(i,x)} \tag{11}$$

Assuming that the upper boundary is confined to the upper half-plane of the SEM image, $\overline{\mathrm{VP}_u}$ is computed for $i \in [0, \mathscr{H}/2]$.

By choosing the valid threshold as half of maximum value, the valid percentage metric for the upper boundary, denoted as VP_u, can be formulated as follows:

$$\mathrm{VP}_u = \frac{\sum_{j=0}^{\mathscr{W}} 1_{\overline{\mathrm{VP}_u} \geq \frac{\max(\overline{\mathrm{VP}_u})}{2}}(j)}{\sum_{j=0}^{\mathscr{W}} 1_{\overline{\mathrm{VP}_u}>0}(j)} \tag{12}$$

where $\mathscr{H}$ denotes the height, $\mathscr{W}$ denotes the width of $\overline{Y}$ respectively. 1 denotes indicator function. The VP_u is between 0 and 1, it represents the percentage of valid vertical gradient pixels under the detected upper contour along the x-axis.

The same metric for the lower boundary VP_l can be obtained by the same way, but change the integration range to $[\mathscr{H}/2, \mathscr{H}]$.

The final metric VP can be obtained from the minimum of VP_u and VP_l:

$$\mathrm{VP} = \min\{\mathrm{VP}_u, \mathrm{VP}_l\} \tag{13}$$

The rationale behind the minimum approach is that artefacts present in a single boundary can already exert a significant influence on the quality of the produced mask. The formulation with the average of VP_u and VP_l was also tested, yet resulted in lower Spearman correlation coefficients than the formulation with minimum.

EXPERIMENTS

BD-MODULE TRAINING

For the boundary discrimination network, we use the Resnet18 (He *et al.*, 2016) network structure. For the training data, each image was manually labelled with the corresponding upper and lower boundary presence labels $y_I^{u,d}$. The total dataset is randomly separated into 5 splits for 5-fold cross validation, and for each split, we take the inference of each trained model on their validation set for subsequent further evaluations.

For data augmentation, each input vertical gradient image was randomly flipped up and down and left and right with a probability of 0.5 (the corresponding labels were inverted accordingly), as well as subjected to a contrast augmentation transform. The loss function employed was the BCE (Binary Cross Entropy) loss function between the predicted labels $\hat{y}_I^{u,d}$ and the true labels $y_I^{u,d}$. For network optimization, the Adam optimizer was used with a learning rate of 1e-5 and L2 regularization coefficient (weight decay) of 5e-2. The batch size was set to 4 due to GPU memory constraint, with the network trained for 200 epochs per split. The model weights with the lowest validation loss were saved and tested.

UNET TRAINING

For the Unet based refinement module, we take the original implementation (Ronneberger *et al.*, 2015) with input channel 1 (SEM image converted to gray scale and resolution 256×256), base feature channel 64 and 5 layers. The pseudo-mask $\hat{M}_I$ resized to resolution 256×256 is taken as label for each input image I. Same as Section 5.1, the performance of the model was evaluated using 5-fold cross-validation on the entire dataset. For optimization, the same optimizer configurations as in Section 5.1 are used, but with batch size 16. For each split, model is trained for 1200 epochs, with early-stopping with patience 200 epochs and min delta 1e-4 used.

EVALUATION METRICS

In order to assess the efficacy of our proposed pipeline, we employed both Intersection Over Union (IoU) and Hausdorff distance as evaluation metrics. The IoU is based on overlapping regions, thus providing a global evaluation. IoU between estimated mask $\hat{Y}$ and ground truth mask Y is defined as follows:

$$IoU(Y,\hat{Y}) = \frac{\left|Y \cap \hat{Y}\right|}{\left|Y \cup \hat{Y}\right|} \tag{14}$$

The Hausdorff distance is based on the maximum distance between two sets:

$$d_{\mathrm{H}}(\hat{Y},Y) := \max\left\{\sup_{\hat{y}\in\hat{Y}} d(\hat{y},Y), \sup_{y\in Y} d(\hat{Y},y)\right\} \tag{15}$$

It is worth noting that all estimated masks are resized back to the original SEM image resolution prior to calculating the metrics, as the Hausdorff distance depends on the resolution.

Table 1. *Quantitative results on the whole dataset (N=429 images). For learning-based methods, the metrics are based on the validation results of the 5-fold cross-validation. SAM denotes pre-trained SAM model with center point as prompt. RS denotes Radon transform (and Sobel operator) based pseudo-mask module. BD denotes Boundary Discrimination module. GT denotes Ground Truth masks.*

Methods	***IoU*↑**	***Hausdorff Distance*↓**
SAM	0.981 (+-0.033)	24.002 (+-50.759)
Unet on GT*	**0.993 (+-0.006)**	16.973 (+-25.659)
RS+BD	0.992 (+-0.006)	**11.447 (+-12.399)**
RS+BD+SAM[a]	0.985 (+-0.006)	18.410 (+-14.632)
Unet on RS+BD[a]	0.987 (+-0.008)	20.709 (+-28.680)

[a]Pseudo-mask based on RS+BD.

*Fully supervised Unet trained with ground-truth mask labels.

RESULTS

This section presents the evaluation of our proposed methods for hair SEM image segmentation, with a focus on both quantitative experimental results and qualitative examples. Furthermore, the discussion will address the proposed auto-evaluation metric, followed by Grad-CAM analysis, which aims at enhancing the interpretability of the BD-module.

BD-MODULE RESULTS

The results of the five-fold cross-validation experiments for the CNN in the BD-module are presented as a weighted F1-score for the upper and lower boundary classes. The weighted F1-score achieved a value of 1.00 across all five splits.

QUANTITATIVE RESULTS

The test results on the whole dataset are presented in Table 1. The *SAM* denotes pre-trained SAM standalone with single center point as prompt[5]. The *RS+BD+SAM* method employs an augmented sampling point with a pseudo-mask based on the *RS+BD* approach, as outlined in Section 4.3.2. *Unet on GT* and *Unet on RS+BD* denote Unet trained with ground truth masks and pseudo-mask from *RS+BD* as labels respectively. Both are with same training conditions in Section 5.2. The results of *Unet on GT* can be used as a reference upper supervised bound for comparison. For the purposes of an ablation study, we separately report the *RS+BD*, *RS+BD+SAM*,*Unet on RS+BD* to compare different refinement methods.

The *Unet on GT* method fully supervised trained with ground truth mask labels achieves optimal test results for IoU, while the improvement compared with *RS+BD* is marginal. It is worth noting that *RS+BD* get best mean and standard deviation for Hausdorff distance compared with other methods, especially significantly outperforming the *SAM* benchmark. This is evidenced by a mean reduction of over 60% in Hausdorff distance and an 80% reduction in standard deviation for both IoU and Hausdorff distance. These results demonstrate the effectiveness of our proposed method based on the Radon transform, Sobel operator and proposed BD-module.

Three following conclusions could be made after analyzing results presented in Table 1.

Pseudo-mask for SAM. A comparison of the *RS+BD+SAM* with the *SAM* benchmark method reveals that the former achieves a superior mean value while significantly reducing the standard deviation with respect to the single center point prompt method. This demonstrates that the use of a pseudo-mask based on the existence and location of the boundary as a priori knowledge through random sampling as prompt improves robustness.

Pseudo-mask for Unet. By comparing *Unet on RS+BD* (Unet trained with pseudo-mask) with *Unet on GT* (Unet trained with ground truth mask), we can observe that the latter gets both better IoU and Hausdorff Distance, which could be explained by the fact that the pseudo-mask labels, which are composed of straight line boundaries, are less precise as supervision signals compared with pixel-level precise ground truth labels, and may provide

[5]The center point constraint is imposed by SEM acquisition condition.

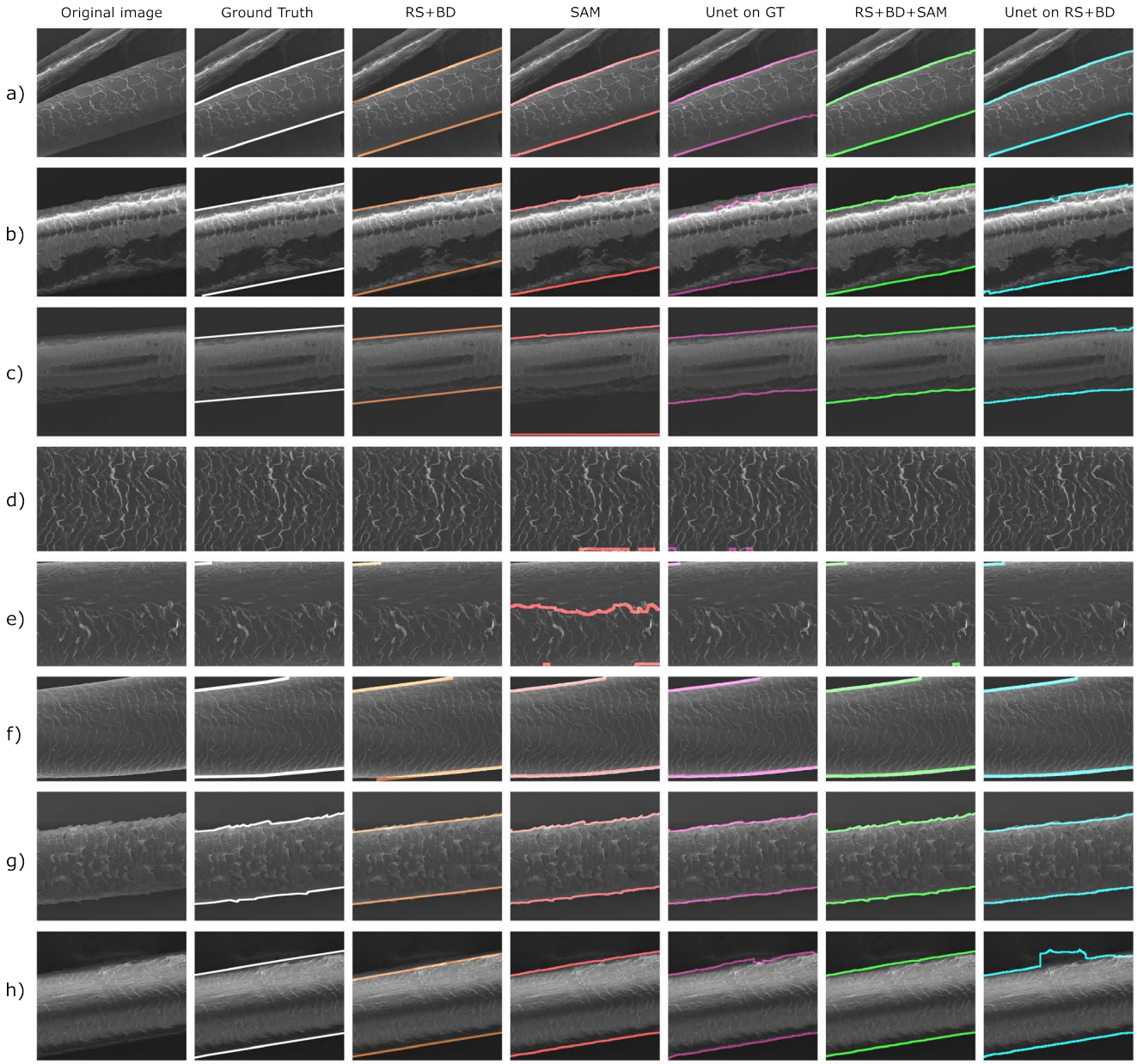

Fig. 4. *Eight qualitative examples (a to h), from left to right: original SEM image, SEM overlapped with ground truth contour (white), contour from RS+BD method (orange), contour from SAM with single point prompt mask (red), contour from Unet trained with ground truth mask labels (magenta), contour from RS+BD with SAM as refinement module (green), contour from Unet on RS+BD as refinement module (cyan).*

erroneous supervision for cases where boundaries are not straight. Whereas the margins between the two methods are tight, the former requires much less annotation (integer existential labels corresponding to upper and lower limits) than the latter (pixel-accurate masks).

Refinement modules. A comparison of the results of the last three rows of Table 1 reveals that the addition of the refinement module (either Unet or SAM) has a negative impact on performance metrics. This is due to the introduction of additional artefacts by both SAM and Unet, which will be illustrated in Section 6.3.

QUALITATIVE RESULTS

We showcase the effectiveness of our proposed method by qualitative examples as shown in Fig 4. The selected examples encompass the principal categories of difficult cases within the dataset: multi-hair (a), low boundary contrast (b, c and h), minor boundary (e), no boundary (d), texture change (e), and non-linear boundary (f and g).

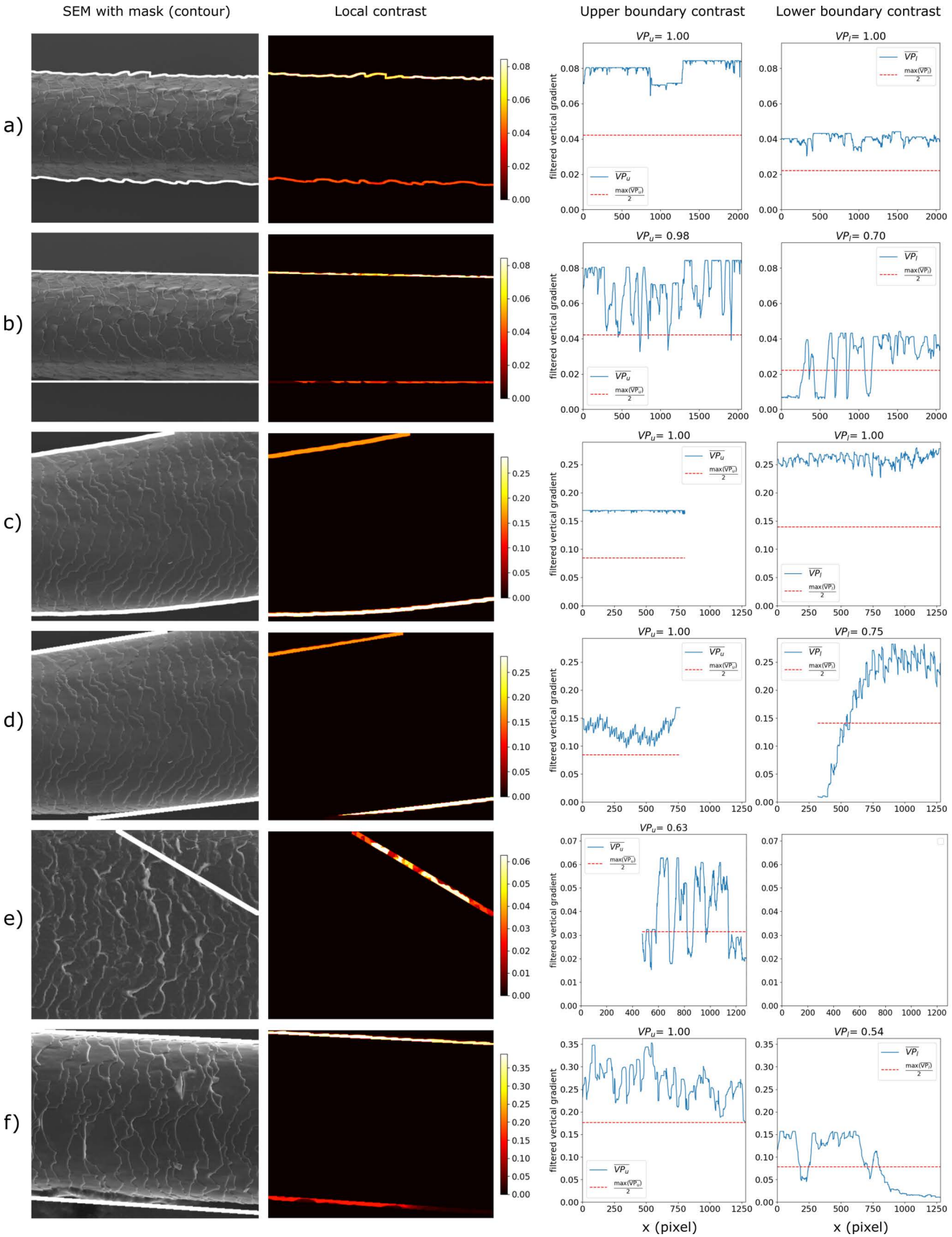

Fig. 5. *Boundary gradient-based valid percentage (VP) calculation process. From left to right: 1. Original SEM image and output mask boundary 2. Vertical gradient under contour 3. Boundary contrast of upper and lower detected boundary gradients, with horizontal coordinates in pixels. The horizontal dot lines represent thresholds for each case (maximum divided by 2). The calculated VP values are in the title of each figure.*

In more precise terms, case a) illustrates the case of multi-hair, due to the constraint of SEM acquisition conditions, when multi-hair occurs at the same time in one image, the one pass center point will be taken as principle one. This is also consistent with the limitation of the search range to the upper and lower half-planes of the sinogram, as described in Section 4.2, and thus the method *RS+BD* is able to successfully segment the main hair.

Examples c) and e) show that the *SAM* method with centroid prompt is prone to artefacts when the texture inside the image changes locally (cuticle and cortex texture region in e), or when the contrast of the same hair boundaries is too low (c). After increasing the number of prompt points by pseudo-mask (method *RS+BD+SAM*), these artefacts were corrected.

The effectiveness of the refinement modules (Unet or SAM) is demonstrated in examples f) and g). In example f), the radius of the hair changes, while in example g), a cuticle warping at the boundary results in an irregular boundary. Consequently, the method (*RS+BD*) that predicted the boundary of the mask to be a straight line both yielded inaccurate results in the lower left corner of f). However, following refinement modules, the predicted mask underwent a further post-hoc correction. However, the refinement modules also introduce new artefacts, as in the example h) in *Unet on RS+BD*, and additional post-processing steps may improve the results.

QUALITY ESTIMATION METRIC RESULTS

Fig. 5 illustrates examples of applying boundary gradient valid percentage (VP) as quality estimation metric. Cases a) and b) represent the same SEM image, but with two different masks derived from *RS+BD* and *SAM* respectively. The first mask exhibits a superior fit to the boundary details (lifted cuticles), thereby yielding higher VPs than the second. The same results are evident in cases c) and d): it is notable that a portion of the lower boundary in d) is erroneous due to the straight line boundary, in contrast, such an error is absent in c), thus resulting in a larger VP for the lower boundary than in d). Cases e) and f) illustrate instances of failed segmentations. In case e), no effective boundary is identified, while in case f), a background artefact in the lower portion is incorrectly detected. As a result, these two cases exhibit comparatively low VP values.

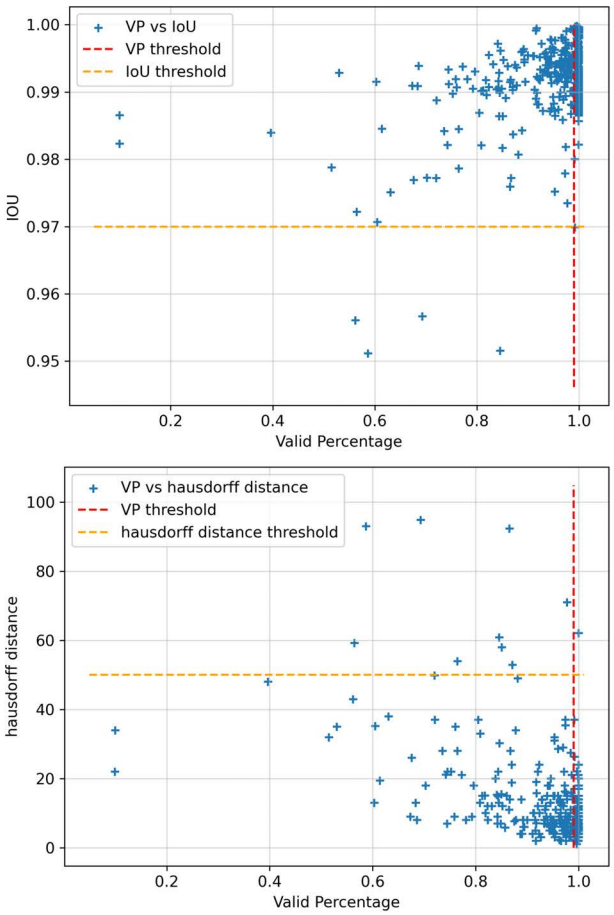

Fig. 6. *Scatter plots of IoU-VP and Hausdorff distance-VP for RS+BD. The VP, IoU and Hausdorff distance thresholds are set to 0.99, 0.97 and 50 respectively for illustration.*

To quantitatively evaluate the VP metric in terms of self-quality estimation on the whole dataset, we show the scatter plots of IoU-VP and Hausdorff distance-VP for method *RS+BD* in Fig. 6. The objective of VP is to accurately identify the produced masks with significant errors. Therefore, the evaluation process is formulated as a classification problem. By setting a threshold for VP, the samples with a VP inferior to the threshold can be estimated as having significant errors (estimated positive). Concurrently, the application of an additional threshold for IoU (or Hausdorff distance) enables the delineation of instances exhibiting *genuinely* substantial errors (true positive).

In accordance with the established convention, the F1 scores, along with the AUC curves, can be depicted in Fig. 7 as a function of varying IoU (or Hausdorff distance) thresholds.

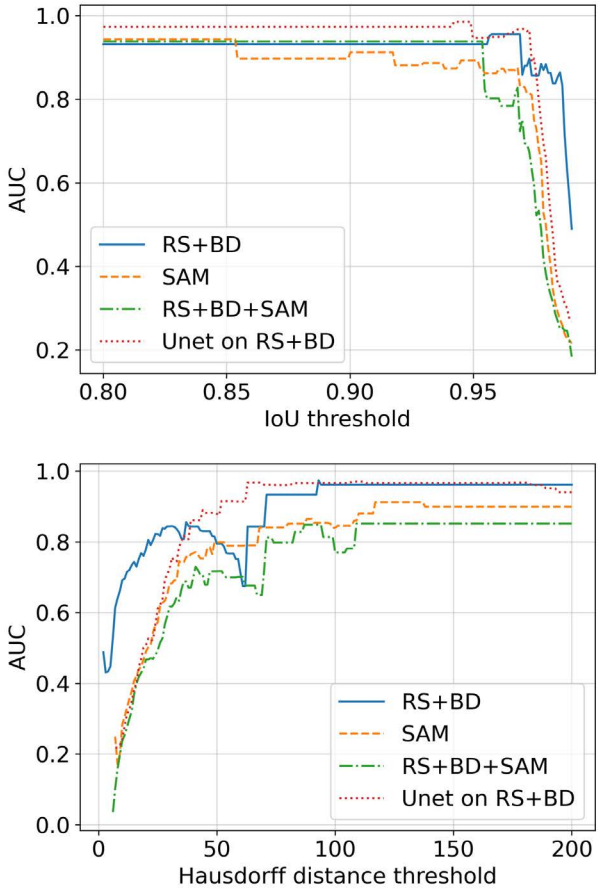

Fig. 7. *AUC curves obtained by varying IoU and Hausdorff distance thresholds, as illustrated in Fig. 6, for different methods.*

For the IoU threshold smaller than 0.95, all the 4 methods get AUC greater or close than 0.9. Especially, *Unet on RS+BD* get AUC close to 0.98. These show the effectiveness of VP in identifying problematic cases when taking IoU as quality indicator. For Hausdorff distance threshold greater than 75, *RS+BD* and *Unet on RS+BD* get AUC around 0.95, but the other two methods get worse AUC.

A comparison of the four methods revealed that the *Unet on RS+BD* method yielded the most favorable AUC values when evaluated on IoU thresholds ≤ 0.95 and Hausdorff distance thresholds ≥ 50 compared with other methods. This demonstrates the high applicability of using VP as a quality assessment metric in this case. It is also noteworthy that for relatively strict thresholds, e.g., IoU thresholds ≥ 0.975 or Hausdorff distance thresholds ≤ 30, *RS+BD* get superior AUC compared with other methods. This can be attributed to the fact that the hairs exhibit predominantly straight boundaries. Furthermore, the ground truth data from the third part algorithm (which represents approximately 80% of the total counts) also displays straight lines, indicating that the results derived from *RS+BD* could potentially be biased. In addition, the results obtained with the *RS+BD* method exhibited fewer irregular artefacts than those derived from the other methods.

Further illustrations of the application of the proposed VP metric in diverse methodologies can be illustrated in Fig. 8.

The notable decline in the VP is attributed to the failure detection of boundaries. Masks with higher VP values demonstrate an enhanced identification of the interface between the background and the hair regions, indicating an improved quality of the mask. For instance, such VP drop can be found in the considerable erroneous mask area resulting from inner cortex texture alterations in cases a) and c), the mask irregularities caused by deep model prediction artefacts in case d), and external errors caused by artefacts in boundaries in cases b), f), g) and h). In the case e) of method SAM, the lower boundary is severely failed, resulting in an extremely low VP (0.36) compared to VPs with other normal error-free masks (close to 1.00). It is noteworthy that for case h), the values of the VPs derived from the four methods are comparatively low in relation to the other cases. This is due to the fact that, as a consequence of the acquisition conditions, the boundaries of the hair in case h) exhibit relatively low contrast. Consequently, the VP metric, which is based on the detected gradient map, is more susceptible to obtaining low values in such conditions.

It is also worth noting that the low VP values observed under such conditions are aligned with the objective of quality estimation: when the boundaries of hair are of sub-optimal quality, the low VP values should be capable of revealing the associated uncertainties. One of the limitations of the VP metric can be demonstrated in the *SAM* method of case i) and the *Unet on RS+BD* method of case f). In both cases, the masks achieve the optimal VP values through the incorporation of well-fit boundary artefacts (adhesive in both instances). However, from a strict standpoint, these artefacts should not be regarded as components of the hairs to be identified. As the VP values only consider gradient information, they are unable to account for semantic aspects.

Consequently, the segmentation quality differences caused by lifted cuticles (which belong to hair) and adhesives (which do not belong to hair) are challenging to distinguish based solely on VP values. Therefore, it would be beneficial for future research to consider the development of a quality metric that also takes semantic meaning into account.

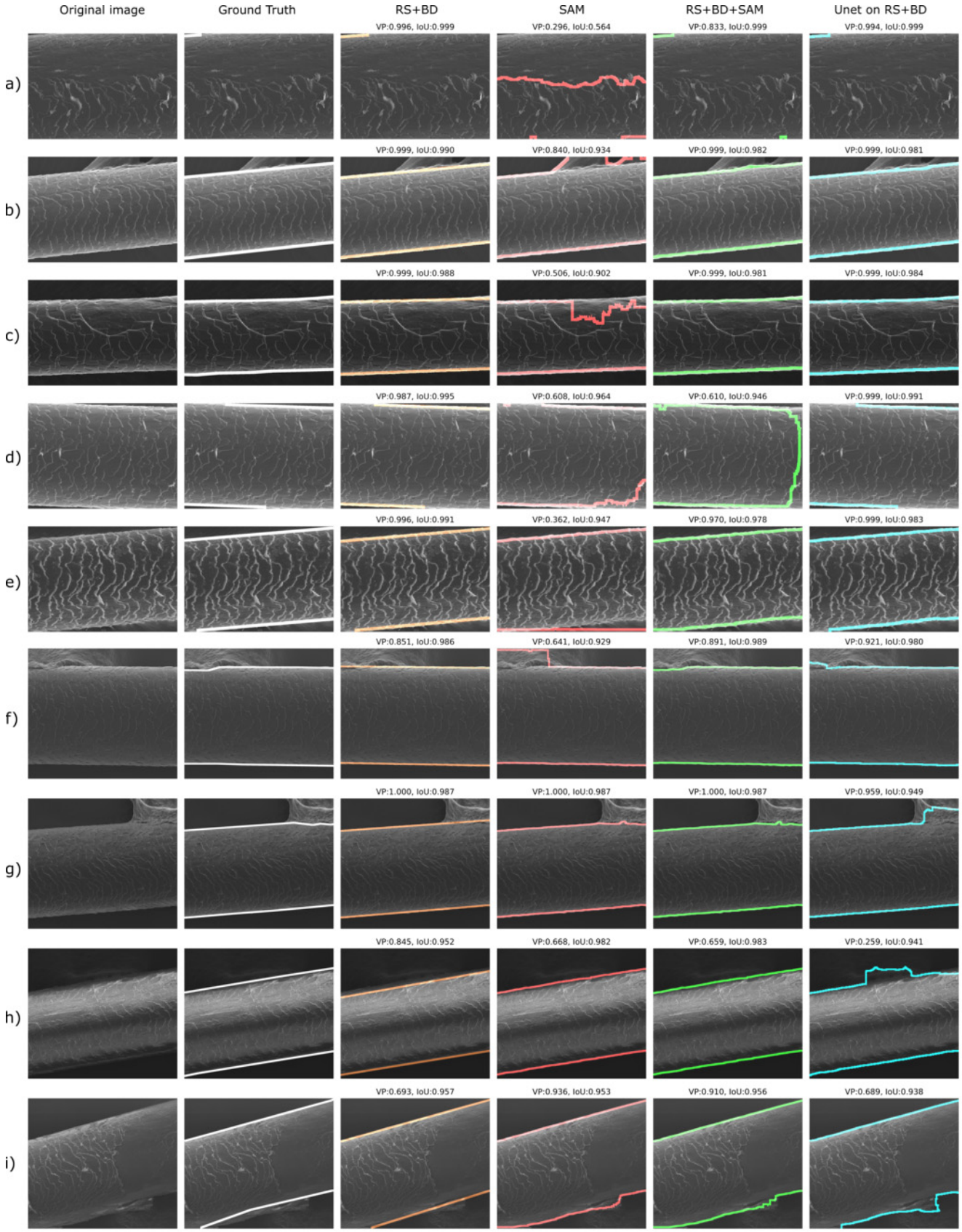

Fig. 8. *Nine segmentation examples with the proposed VP (valid percentage) auto-evaluation metric, which are shown in the title for each image. IoUs are calculated with the ground truth masks. The colors of the contours are the same as in Fig4. The examples are obtained with structuring element size W=20 (cf. Section. 4.4).*

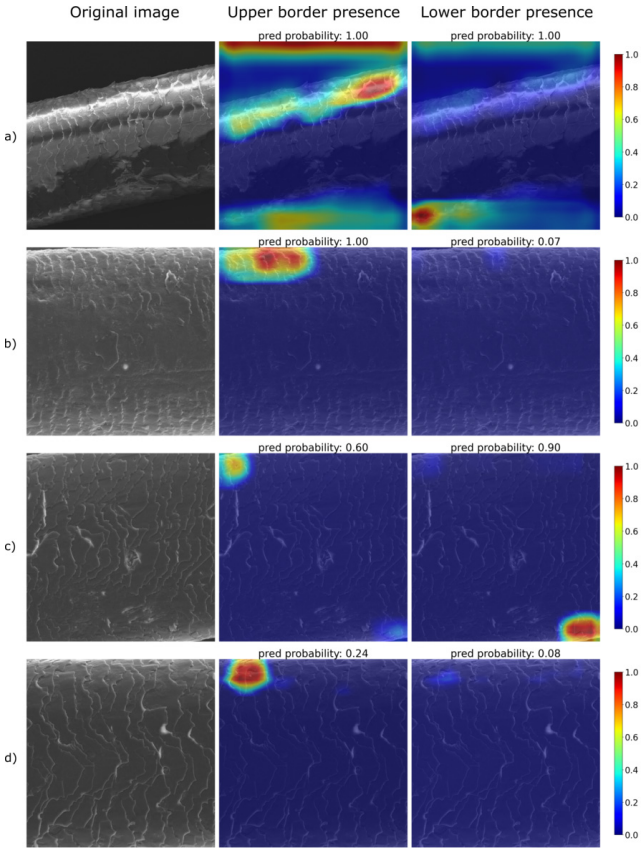

Fig. 9. *Grad-CAM examples on validation set for trained CNN, from left to right: original image, saliency map for upper boundary, saliency map for lower boundary. The predicted probability for existence of upper/lower boundaries are indicated at title for each CAM image.*

GRAD-CAM ON BD-MODULE

We further illustrate the effectiveness and improve the interpretability of BD-module by means of Grad-CAM (Selvaraju *et al.*, 2017) results of the trained CNN on validation set, as shown in Fig. 9.

In Fig.9, example a) illustrates a scenario where the SEM image comprises two boundaries and the SEM acquisition low-light condition at the same time. In this instance, the CNN's attention for the presence of the upper and lower boundaries is distributed between the boundaries themselves and the background. This aligns with human perception. However, the activation regions in the lower part of the upper border presence column and in the upper part of the lower border presence column are not fully expected.

Examples b) and c) illustrate cases with a single boundary or partial boundaries. In these instances, the focus is concentrated at the boundary locations, which ultimately yields probability values that lead to the correct results.

It is noteworthy that in case d), the upper boundary is not visible, yet in the corresponding plot, the network attention is still focused on the region of the plot where this boundary is most likely to be present.

The presented examples in Fig.9 demonstrate that the CNN network has been effectively trained and enhance the interpretability of the proposed BD-module.

CONCLUSION

This paper presents a novel weakly supervised pipeline for hair SEM image segmentation. The proposed method is based on the Radon transform and the Sobel operator, with an integrated CNN module to estimate the presence/absence of upper and lower boundaries, and a refinement module based on Unet or SAM to further cope with cases where boundaries have significant changes in diameter.

We also benchmark our SEM hair dataset with other methods such as Unet and SAM. Experimental results show that although our proposed method uses only image-level annotations, it achieves comparable performance on IoU, as well as more than 30% improvement in mean and more than 50% in standard deviation on Hausdorff distance compared to Unet trained in a fully supervised way. Furthermore, ablation studies show the effectiveness of each proposed component.

Furthermore, our proposed pipeline, which is mainly based on classical image processing techniques, has better interpretability than end-to-end deep learning based methods such as Unet and SAM, with the Grad-CAM results improving the interpretability for the BD-module.

For automatic quality evaluation and to improve the applicability of our proposed methods to other SEM microscopic hair images, we further propose a quality estimation metric based on detected boundary gradient map (VP). The quantitative results and qualitative examples demonstrate the effectiveness of the proposed metric in identifying instances of significant segmentation error, thereby enabling the automatic assessment of segmentation quality for large-scale datasets.

One of the limitations of this work is the refinement modules, as deep learning-based refinement modules (Unet or SAM) simultaneously refine the boundaries and introduce new artefacts. Therefore, exploring alternative refinement methods could be promising avenues for future research. Another limitation is the proposed VP metric, which

currently only uses the homogeneity of the detected boundary gradient, but doesn't take into account the semantic information, which could be useful for accurately identifying boundary details and eliminating artefacts.

The SEM hair microscopy dataset of 429 images will be made open source to further benefit the research community.

CONFLICTS OF INTEREST

The authors declare that there are no known conflicts of interest associated with this publication. Thérèse Baldeweck, Pierre Dupuis and Thomas Bornschloegl are full employees of L'Oréal Research and Innovation.

REFERENCES

Ahn J, Kwak S (2018). Learning pixel-level semantic affinity with image-level supervision for weakly supervised semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition.

Angulo J, Velasco-Forero S, Chanussot J (2009). Multiscale stochastic watershed for unsupervised hyperspectral image segmentation. In: 2009 IEEE international geoscience and remote sensing symposium, vol. 3. IEEE.

Bhaskar H, Werghi N (2011). Comparing Hough and Radon transform based parallelogram detection. In: 2011 IEEE GCC Conference and Exhibition (GCC). IEEE.

Bilodeau A, Delmas CV, Parent M, De Koninck P, Durand A, Lavoie-Cardinal F (2022). Microscopy analysis neural network to solve detection, enumeration and segmentation from image-level annotations. Nat Mach Intell 4:455–66.

Chen H, Wang J, Chen HC, Zhen X, Zheng F, Ji R, Shao L (2021). Seminar learning for click-level weakly supervised semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision.

Chu F, Anex DS, Jones AD, Hart BR (2020). Automated analysis of scanning electron microscopic images for assessment of hair surface damage. R Soc Open Sci 7:191438.

Coroaba A, Chiriac AE, Sacarescu L, Pinteala T, Minea B, Ibanescu SA, Pertea M, Moraru A, Esanu I, Maier SS, *et al.* (2020). New insights into human hair: SAXS, SEM, TEM and EDX for Alopecia Areata investigations. PeerJ 8:e8376.

Dong M, Liu D, Xiong Z, Chen X, Zhang Y, Zha ZJ, Bi G, Wu F (2020). Towards neuron segmentation from macaque brain images: a weakly supervised approach. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23. Springer.

Dutta A, Zisserman A (2019). The VIA annotation software for images, audio and video. In: Proceedings of the 27th ACM International Conference on Multimedia, MM '19. New York, NY, USA: ACM.

Galliano A, Ye C, Su F, Wang C, Rakshit R, Guerin M, Flament F, Steel A (2023). Assessing the effect of cleansing products on artificially polluted human hairs and skin through in vivo and in vitro models. Skin Res Technol 29:e13220.

Galliano A, Ye C, Su F, Wang C, Wang Y, Liu C, Wagle A, Guerin M, Flament F, Steel A (2017). Particulate matter adheres to human hair exposed to severe aerial pollution: consequences for certain hair surface properties. Int J Cosmet Sci 39:610–6.

He K, Zhang X, Ren S, Sun J (2016). Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition.

Huang Z, Wang X, Wang J, Liu W, Wang J (2018). Weakly-supervised semantic segmentation network with deep seeded region growing. In: Proceedings of the IEEE conference on computer vision and pattern recognition.

Kim YD, Jeon SY, Ji JH, Lee WS (2010). Development of a classification system for extrinsic hair damage: Standard grading of electron microscopic findings of damaged hairs. Am J Dermatopathol 32:432–8.

Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo WY, *et al.* (2023). Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision.

Kniesel H, Sick L, Payer T, Bergner T, Devan KS, Read C, Walther P, Ropinski T, Hermosilla P (2023). Weakly supervised virus capsid detection with image-level annotations in electron microscopy images. In: The Twelfth International Conference on Learning Representations.

Kulharia V, Chandra S, Agrawal A, Torr P, Tyagi A (2020). Box2seg: Attention weighted loss and discriminative feature learning for weakly supervised segmentation. In: European Conference on Computer Vision. Springer.

Lee H, Jeong WK (2020). Scribble2label: Scribble-supervised cell segmentation via self-generating pseudo-labels with consistency. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23. Springer.

Lee I, Du X, Anthony B (2017). Hair segmentation using adaptive threshold from edge and branch length measures. Comput Biol Med 89:314–4.

Lee S, Choi A, Baek J, Kim H, Shin M, Koh J (2016). Twelve-point scale grading system of scanning electron microscopic examination to investigate subtle changes in damaged hair surface. Skin Res Technol 22:406–11.

Li J, Fan J, Zhang Z (2022). Towards noiseless object contours for weakly supervised semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

Li S, Ren M, Ach T, Gerig G (2023). Microscopy image segmentation via point and shape regularized data synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer.

Liu X, Blusseau S, Velasco-Forero S (2024). Counting melanocytes with trainable h-maxima and connected component layers. In: International Conference on Discrete Geometry and Mathematical Morphology. Springer.

Man Q, Zhang L, Cho Y (2021). Efficient hair damage detection using SEM images based on convolutional neural network. Appl Sci 11:7333.

Martins da Cruz JM, Sangalli M, Decencière E, Velasco-Forero S, Baldeweck T (2024). A posteriori deep learning segmentation quality estimation based on prediction entropy. Image Anal Stereol 53:121–30.

Obikane S, Aoki Y (2020). Weakly supervised domain adaptation with point supervision in histopathological image segmentation. In: Pattern Recognition: ACPR 2019 Workshops, Auckland, New Zealand, November 26, 2019, Proceedings 5. Springer.

Qiu D, Xiong S, Yi J, Peng J (2024). Weakly-supervised cross-domain segmentation of electron microscopy with sparse point annotation. IEEE Transactions on Big Data .

Rivest JF, Soille P, Beucher S (1992). Morphological gradients. In: Nonlinear image processing III, vol. 1658. SPIE.

Ronneberger O, Fischer P, Brox T (2015). U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer.

Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision.

Shih HC (2014). An unsupervised hair segmentation and counting system in microscopy images. IEEE Sensors Journal 15:3565–72.

Tomes C, Jones J, Carr C, Jones D (2007). Three-dimensional imaging and analysis of the surface of hair fibres using scanning electron microscopy. Int J Cosmet Sci 29:293–9.

van Ginkel M, Hendriks CL, van Vliet LJ (2004). A short introduction to the radon and hough transforms and how they relate to each other. Delft University of Technology .

Vincent L (1993). Grayscale area openings and closings, their efficient implementation and applications. In: First Workshop on Mathematical Morphology and its Applications to Signal Processing. Citeseer.

Wang Z, Wang E, Zhu Y (2020). Image segmentation evaluation: a survey of methods. Artif Intell Rev 53:5637–74.

Zhang L, Man Q, Cho YI (2021). Deep-learning-based hair damage diagnosis method applying scanning electron microscopy images. Diagnostics 11:1831.

Zhao T, Yin Z (2020). Weakly supervised cell segmentation by point annotation. IEEE Transactions on Medical Imaging 40:2736–47.