IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 9, 2021, accepted December 22, 2021, date of publication December 31, 2021,
date of current version January 13, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3140031

Skin Microstructure Segmentation and Aging Classification Using CNN-Based Models

CHO-I MOON[1] AND ONSEOK LEE[1,2]

[1]Department of Software Convergence, Graduate School, Soonchunhyang University, Asan-si, Chungcheongnam-do 31538, Republic of Korea
[2]Department of Medical IT Engineering, College of Medical Sciences, Soonchunhyang University, Asan-si, Chungcheongnam-do 31538, Republic of Korea

Corresponding author: Onseok Lee (leeos@sch.ac.kr)

This work was supported in part by the Soonchunhyang University Research Fund, in part by the Korea Government [Ministry of Science and ICT (MSIT)] through the National Research Foundation of Korea (NRF) under Grant 2019R1F1A1058827, and in part by the Ministry of Education under Grant 2020R1A6A3A13072938.

ABSTRACT The skin surface is composed of a network-like microstructure comprising wrinkles. Observing and analyzing the microstructure of the skin that changes with the skin condition and aging are simple, stable, and accurate evaluation methods for skin diagnosis. However, the skin surface includes various morphological and topological changes, depending on the individual or the degree of aging. It is difficult to accurately extract and analyze a skin microstructure including these changes. Therefore, we perform skin microstructure segmentation and aging analysis by using convolutional neural network (CNN) models. First, we propose a fusion UNet model to extract the skin microstructure. We compare and evaluate the segmentation performance by using an image processing method and deep learning models. Next, we classify skin aging based on the skin microstructure. For the classification, we use four mobile CNN models: NASNet-Mobile, MobileNetV2, MobileNetV3-Small, and EfficientNet-B0. Subsequently, we compare and evaluate their classification performances. Results show that the segmentation images of the fusion U-Net are most similar to the ground truth, and the fusion U-Net model can detect fine wrinkles that are difficult to identify by the naked eye. In the microstructure-based classification of skin aging, MobileNetV3-Small exhibits the best performance with an accuracy of 94%. The proposed method facilitates an objective and quantitative analysis of the skin surface with more diverse aging characteristics. Consequently, the association between skin aging and skin microstructure changes is confirmed. Our study can be utilized in the diagnostic studies on various skin characteristics, including skin texture, anisotropy, and roughness. The proposed method can also be applied to a mobile-based self-diagnosis system.

INDEX TERMS Skin aging, skin micro-structure segmentation, aging classification, CNN models, skin condition evaluation.

I. INTRODUCTION

The skin surface comprises a network of polygonal structures composed of wrinkles. The network of the skin surface changes over time as a result of extrinsic factors such as ultraviolet rays and the surrounding environment and intrinsic factors such as aging. The skin surface represents the morphological changes in wrinkles and wrinkled cells, and the topological changes in epidermal become more noticeable as the elastic fibers of the papillary dermis loosen or disappear because of aging [1]. Analyzing and tracking the morphological and topological changes in the skin are crucial for estimating the current skin condition or skin aging.

The associate editor coordinating the review of this manuscript and approving it for publication was Gulistan Raja.

Numerous studies have investigated the skin surface microstructure and verified the relationship between skin characteristic changes and aging. In some studies, skin conditions have been analyzed and evaluated by determining the length and width of wrinkles and area of wrinkled cells from the skin surface structure [2], [3]. In other studies, skin structure changes with age have been estimated by analyzing the association between skin aging and factors such as lifestyle, age, and gender [4]. In most skin-related studies, a skin diagnosis device, such as a dermoscope, is used to obtain a formalized skin image [5]–[7]. However, a specialized equipment is challenging to use and expensive. Therefore, in our previous study, we acquired unconstrained skin images using a mobile device and conducted a skin microstructure analysis of the mobile images [8]–[11]. However, the analysis of the

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

skin microstructure was difficult for topological changes such as blood vessels appearing prominently on the skin surface in the mobile images.

Recently, deep convolutional neural networks (CNNs) have exhibited excellent performance in various computer vision studies. Deep learning models have been used in segmentation and classification tasks on various medical images requiring subjective interpretations by experts [12]–[14]. In several skin-related studies involving deep learning, the segmentation or classification of skin diseases, such as melanoma, has been actively performed [15]–[17]. In addition, in certain studies, automatic age classification has been performed by using two-dimensional facial images [18]. Age estimation and classification are complicated as they are affected by various factors, such as human lifestyle, health, and gender. Nevertheless, improved models for age classification have been proposed in many studies. To date, classification studies on aging have utilized only facial images with various age-related features. We can directly observe skin components, such as wrinkles and wrinkled cells in the skin surface microstructure. Therefore, we segmented the skin microstructure and classified skin aging based on the morphological features of the structure using CNN-based models.

We used an unconstrained mobile skin image that was exposed to surrounding environments, such as illumination and shadows, and involved challenging factors such as the skin curvature. Also, the mobile images include the difficulties of the skin itself: the skin surface exhibited prominent topological characteristics with aging, and wrinkles in young skin were difficult to recognize by the naked eye owing to their fineness. To overcome these challenges, we propose a fusion UNet model as a segmentation method for the skin surface microstructure. We compared and analyzed the segmentation performance using the previous image processing method and recent deep learning models, such as fully convolutional network (FCN-8) and ResUNet++. Subsequently, we classified skin aging conditions using segmented skin structure images. As age classification, we used four different CNN models: NASNet-Mobile, MobileNetV2, MobileNetV3-Small, and EfficientNet-B0. We compared and analyzed the classification performance of these four mobile models. In addition, the optimum method was presented for evaluating skin aging in a mobile environment.

Our contributions are as follows.

1) We conducted a study by employing deep learning models that operate robustly on unstructured mobile images. Skin images of the back of a hand were acquired using a mobile camera. The skin on the back of the hand showed morphological changes and topological changes with aging. The mobile skin image was an unconstrained image that contains skin aging features and various noises.
2) We proposed a fusion UNet model to extract the skin microstructure. It is difficult to accurately identify the skin microstructure in mobile skin images owing to various factors, such as hair, lighting, and shadows. Therefore, we used the proposed model to extract the skin microstructure including the details difficult to identify with the naked eye.
3) We used four mobile models to classify and analyze skin aging. MobileNetV3-Small was the lightest and exhibited the best classification performance with an accuracy of 94%. We confirmed that it is possible to evaluate skin aging in a mobile environment.

II. SKIN SURFACE IMAGE DATASET

We used a Galaxy A3 smartphone (SM-A310NZKAKOO, Samsung Electronics Co., Republic of Korea) to capture the images of the skin on the back of a hand. The skin on the back of the hand can be observed with the naked eye to reveal the skin microstructure and changes related to aging. The subjects were 68 healthy Cambodians, who were non-smokers, in the age range of 10–80 years. The mobile skin images were acquired by zooming 2.0× using a rear camera (resolution 1.3 MP; aperture F1.9). When acquiring images, we photographed so that the area of the back of the hand was visible as much as possible in the image. We did not set any restrictions on the surrounding environment to deal with unstructured images that the anyone can acquire in daily life. So, we only did the camera software settings.

The acquired images were unconstrained images exposed to the surrounding environment and various noises. Figure 1 shows the acquired mobile images, which contain challenging factors that make it difficult to observe the skin microstructure, for example, prominent curvatures in blood vessels (first image), reflections or shadows (second and third images), and hair or scars (fourth image). The size of the acquired original image is 2064 × 1161 pixels. We cropped the original image to a size of 516 × 387 pixels, selected the cropped images including the skin on the back of the hand for post-processing. Finally, 8-10 crop images were selected from one subject's image, and the final selected image resized 256 × 256 pixels. We constructed the final skin image dataset.

First, for the segmentation of the skin microstructure, we used 45 different images to train and applied data augmentation such as rotation, shift, shear, zoom, and flip. The microstructure of the skin on the back of the hand can be identified with the naked eye. Therefore, biomedical engineer researchers, who had background knowledge on skin, generated the ground truth (GT) of the skin microstructure that was confirmed with the naked eye. The GT was used as the criterion for evaluating the segmentation performance.

Next, we evaluated the skin conditions in the mobile images by the naked eye and categorized them into three groups: young, middle, and old. The average ages of the subjects in the young, middle, and old groups were 23, 39, and 60 years, respectively. The training dataset for classification was applied with data augmentation (rotation, left and right

FIGURE 1. Skin images were acquired using a mobile device. Mobile skin images contain morphological and topological changes related to aging and various environmental noises.

TABLE 1. Mobile skin image datasets of aging group.

Group	The number of original images	The number of crop images	The number of augmentation images (Train dataset)
Young	21	186	1,490
Middle	22	208	1,660
Old	25	227	1,820

flips, scale, and shift). Table 1 lists the number of images in the three groups used for aging classification. To evaluate the classification performance, we split the images of each group into training and test datasets at a ratio of 8:2. We used 120 images for testing and did not apply additional data augmentation to the test dataset. In this study, researchers with experience in skin aging studies directly evaluated the results.

III. METHODS

We used CNN models for segmentation and classification using GeForce RTX 2080 Ti 11 GB × 2. We used the TensorFlow framework in our experiments.

A. SKIN MICROSTRUCTURE SEGMENTATION

The U-Net structure is based on a fully convolutional network and adds the skip connections that connect the encoding and decoding features corresponding to each other. Therefore, it is a simple and effective model for medical image segmentation [19], [20]. We added a fusion module that concatenated the upsampling layers of the decoding process in the basic U-Net structure. The fusion operation complemented the image details. Therefore, to improve the performance of detecting fine wrinkles constituting the skin structure, we segmented the skin microstructure using the fusion U-Net model (Figure 2). For model training, we set the batch size to 2, epoch to 200 times, loss function to cross entropy, and optimization function to Adam.

To evaluate the segmentation performance, we compared the GT with the five segmentation results of the watershed-algorithm-based segmentation method of a previous study [8], FCN-8, UNet, ResUNet++, and the proposed fusion U-Net. The FCN is an end-to-end model that consists of only convolution layers. We used a model based on VGG-16 [21]. ResUNet++ uses the residual block, squeeze and excitation block, atrous spatial pyramidal pooling, and attention block to improve segmentation performance; it works well for small data [22]. We analyzed and compared the segmentation results using the image processing method and the CNN-based segmentation models. Then, we presented the optimal method for skin microstructure segmentation. We used five similarity indicators to evaluate the segmentation performance: Dice similarity coefficient (DSC) [23], Jaccard similarity coefficient (JSC) [24], modified Hausdorff distance (MHD) [25], adjusted Rand index (ARI) [26], and variation of information (VI) [27]. The similarity between two images, A and B, was calculated in terms of each evaluation indicator as follows: the higher the similarity between A and B, the larger were the DSC, JSC, and ARI values, and inversely, the smaller were the MHD and VI values.

$$DSC = \frac{2\,|A \cap B|}{|A| \cup |B|} \tag{1}$$

$$JSC = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

$$MHD = max\left\{h\left(A,B\right), h\left(B,A\right)\right\},$$
$$h\left(A,B\right) = \sum\nolimits_{A_i}\left(\min_{B_j} d(A_i - B_j)\right) \tag{3}$$

$$ARI(A,B) = \frac{\sum_{i,j}\binom{n_{ij}}{2} - \frac{\left[\sum_i \binom{a_i}{2}\sum_i \binom{b_i}{2}\right]}{\binom{n}{2}}}{\frac{1}{2}\left(\sum_i \binom{a_i}{2} + \sum_i \binom{b_i}{2}\right) - \frac{\left[\sum_i \binom{a_i}{2}\sum_i \binom{b_i}{2}\right]}{\binom{n}{2}}} \tag{4}$$

$$VI\left(A,B\right) = -\sum_{i,j} r_{ij}\left[log\left(\frac{r_{ij}}{p_i}\right) + log\left(\frac{r_{ij}}{q_i}\right)\right],$$
$$p_i = \frac{|A_i|}{n} \quad and \; q_i = \frac{|B_i|}{n}, \; r_{ij} = |A_i \cap B_i| / n \tag{5}$$

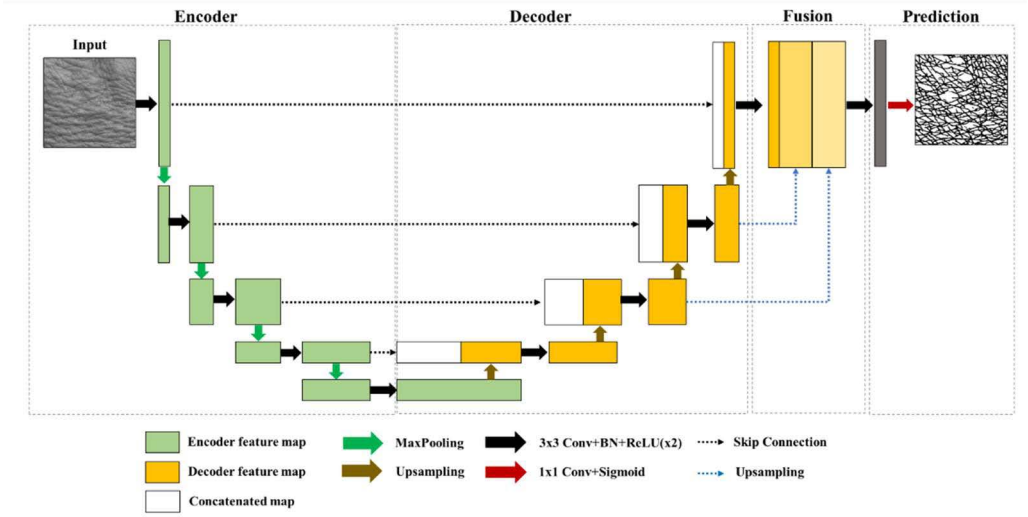

FIGURE 2. Structure of the fusion U-Net. We added a fusion module to the U-Net structure to improve the segmentation performance of extracting the details of the skin microstructure.

TABLE 2. Detailed information of mobile DCNNs.

Models	Params (M)	Flops (B)
NASNet-Mobile	5.6	9.19
MobileNetV2	3.8	6.43
MobileNetV3-Small	2.8	1.84
EfficientNet-B0	5.6	6.38

B. SKIN AGING CLASSIFICATION

We classified aging using mobile models to analyze skin aging in a mobile environment. We used NASNet-Mobile, MobileNetV2, MobileNetV3-Small, and EfficientNet-B0 (Table 2). We performed transfer learning using four models pre-trained on ImageNet.

NASNet is a model that determines the optimal structure using reinforcement learning. It constructs the optimal convolution cells using an RNN controller that determines the stride and filter size of the convolution layer and stacks them several times. It is scalable to large datasets and exhibits a high performance [28].

MobileNetV2 reduces the model computation using an inverted residual block that performs the skip connection of the feature map of the low-dimensional layers. It improves the model performance by preventing information loss through a linear bottleneck layer [29].

MobileNetV3-Small is a model that explores the optimized structure using the neural architecture search and NetAdapt algorithm. It introduces a new nonlinear activation function, named h-swish, reduces latency, and improves accuracy over previous models [30].

EfficientNet-B0 is a model that efficiently adjusts the depth, channel width, and resolution of the model using the compounding scaling method and determines the optimal combination. Thus, it performs well with much fewer parameters [31].

We set the batch size to 2, epoch to 100 times, loss function to cross entropy, and optimization function to Adam to train classifiers. In addition, we used accuracy, recall, precision, and F1-score to evaluate the performances of the four classifiers.

IV. RESULTS

A. SKIN MICROSTRUCTURE SEGMENTATION BASED ON THE FUSION UNET MODEL

In the aging skin (old group), the dermis and epidermis separate, blood vessels become prominent, and severe curvature occurs (first-row image of Figure 3). The skin in the process of aging (middle group) has ambiguous wrinkles during deepening or disappearance (second and third row images of Figure 3, respectively). The young skin (young group) has very complex and dense wrinkles (fourth row image of Figure 3). In addition, the mobile skin image includes various noises such as hair, scars, blurring, and shadows on the skin surface (fifth row image of Figure 3). Different characteristics with skin aging and various external noises make it difficult to analyze the skin. However, a precise observation of the skin surface is required to evaluate the skin condition by analyzing the shapes and patterns of the skin microstructure. It is difficult to observe the fine wrinkles of skin only with the naked eye, without using special equipment such as a dermoscope. Therefore, it is necessary to perform the segmentation of the skin microstructure to increase the accuracy of skin aging evaluation based on mobile skin images.

Figure 3 shows the results of skin microstructure segmentation. The existing watershed-algorithm-based segmentation method always performed segmentation in the form

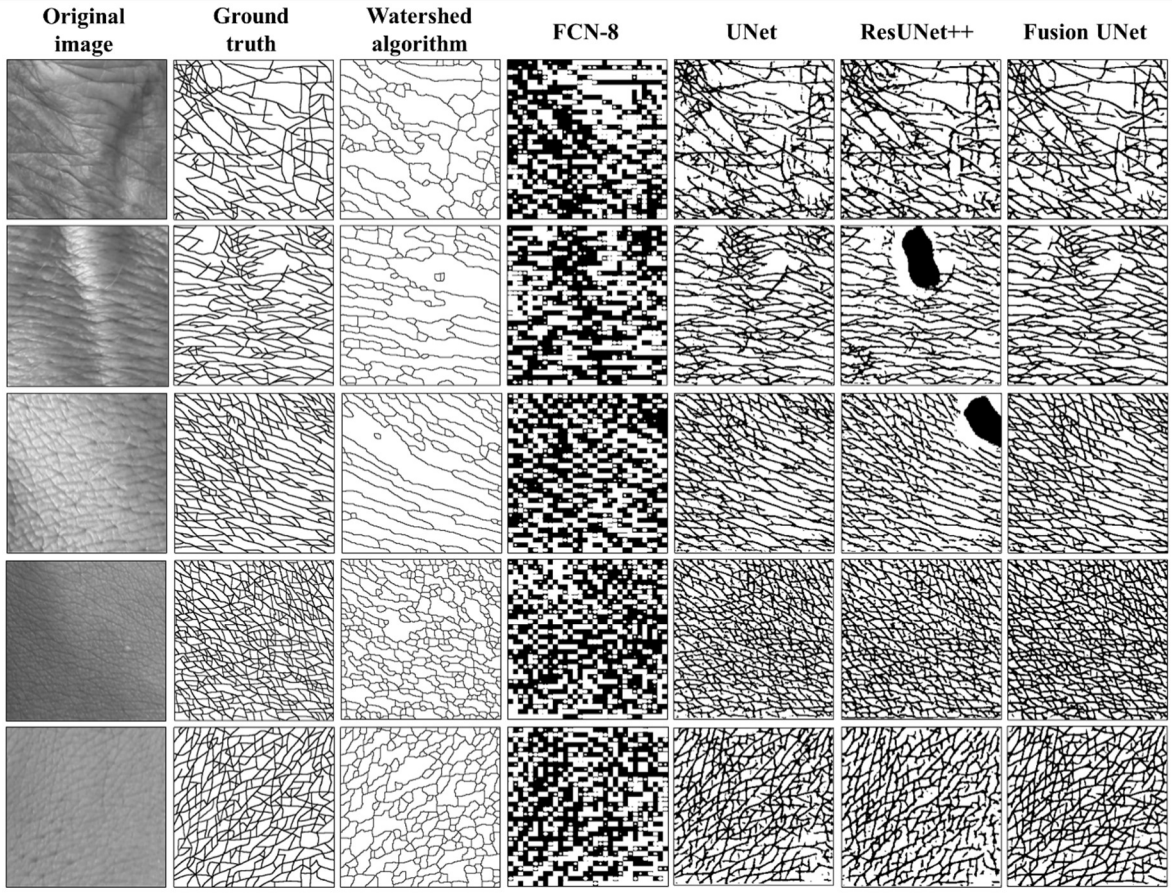

FIGURE 3. Results of skin microstructure segmentation. The figure shows the original images of skin (first column), GT (second column), results of the watershed-algorithm-based segmentation method (third column), results of FCN-8, UNet, and ResUNet++ models (fourth, fifth, and sixth columns, respectively), and results of the proposed fusion U-Net model (seventh column).

of closed wrinkles based on the wrinkled cell area. Therefore, the watershed-algorithm-based segmentation method could not detect wrinkles around connected wrinkled cells and broken wrinkles during creation or disappearance. These problems resulted in incorrect segmentation areas that recognized multiple wrinkled cell regions as one (the second and third column image). This method was also susceptible to noise such as pores and hair on the skin surface, resulting in over-segmentation. The results of FCN-8 showed that the estimated wrinkle area was thick, and the representation of wrinkles was very rough. The results of UNet were similar to the GT, but they contained considerable noise around wrinkles. In addition, the results of ResUNet++ were significantly affected by the illumination or shadow. Mis-segmentation occurred in some areas with significant shading (the second and third column images). Finally, the results of the fusion UNet were similar to the GT, and they had less noise regions than those of UNet and ResUNet++. In addition, the fusion UNet detected small and fine wrinkles that were difficult to accurately recognize with the naked eye. Therefore, we can analyze the directionality or anisotropy of the skin structure using the fusion UNet results.

Next, we evaluated the similarity between the GT and the results of the five segmentation methods. Table 3 presents the results of the similarity indicator. All five indicators showed that the segmentation result of the fusion UNet was the most similar to the GT. The result of FCN-8 was the worst. UNet and ResUNet++ models achieved low-similarity results owing to noises and over-segmentation of the fine wrinkles. The overlap area between the segmentation results of the fusion UNet and the GT was large (the results of the DSC, JSC, ARI, and VI), and the spatial positions of wrinkles in two images were similar (the results of the MHD). Therefore, we confirmed that the proposed model, fusion UNet, extracted the microstructure of the skin considering topological and morphological features with aging.

B. SKIN AGING CLASSIFICATION USING THE SKIN MICROSTRUCTURE

We divided skin microstructure images into three groups, young, middle, and old, according to the skin conditions, and then constructed the classification dataset. Figure 4 shows the microstructure segmentation images of the skin using

TABLE 3. Results of five similarity indicators.

Methods	DSC	JSC	MHD	ARI	VI
Watershed-algorithm-based segmentation	0.85	0.74	6.95	0.57	1.85
FCN-8	0.68	0.51	9.20	0.53	1.78
UNet	0.88	0.78	6.57	0.72	1.74
ResUNet++	0.87	0.78	6.65	0.71	1.77
fusion UNet	**0.89**	**0.80**	**6.20**	**0.74**	**1.65**

TABLE 4. Results of skin aging classification performance for four CNN-based classifiers.

Model	Accuracy	Recall	Precision	F1-score
NASNet-Mobile	0.92	0.92	0.92	0.92
MobileNetV2	0.93	0.93	0.92	0.92
MobileNetV3-Small	**0.94**	**0.94**	**0.94**	**0.94**
EfficientNet-B0	0.93	0.94	0.93	0.93

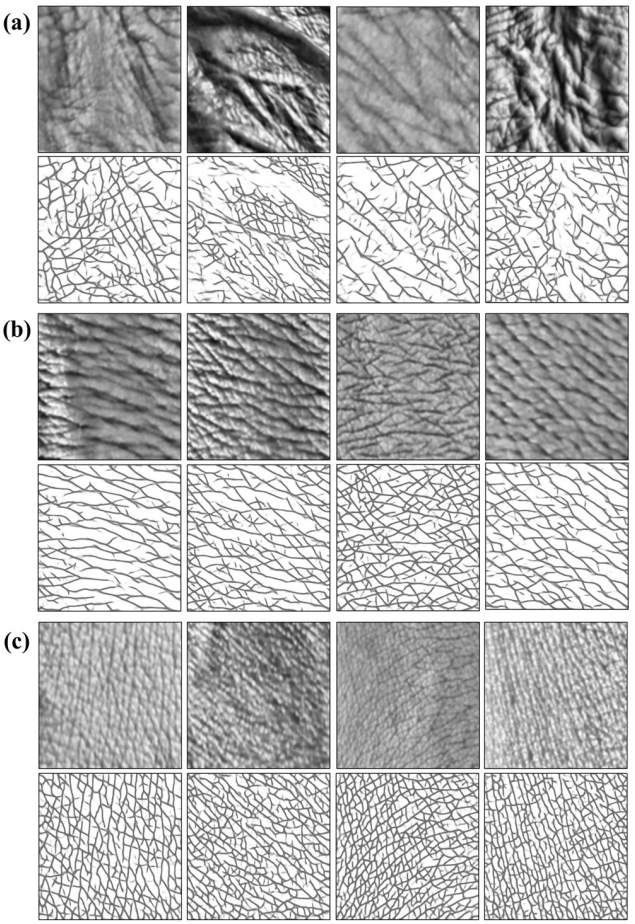

FIGURE 4. Results of the skin microstructure segmentation of three group images [(a) are the old group images, (b) are the middle group images, and (c) are the young group images] using the fusion U-Net model.

the fusion U-Net. The skin of the young group comprises a very dense and complex polygonal structure surrounded by fine wrinkles. The skin surface made of these irregular structures is tight. In the middle group, small and fine wrinkles of the skin disappear, and a skin structure with primary wrinkles is formed. This structure has a particular pattern and directionality. In the old group, skin components such as wrinkles disappear, and the shapes of the skin structure is lost. Further, deep wrinkles are prominent, and in particular, strong curvature appears on the skin surface owing to the blood vessels. We classified skin aging using microstructure images containing aging features. We used four CNN models and evaluated their classification performances.

We evaluated the classification performance using a test set. Table 4 presents the results of the four indicators (accuracy, recall, precision, and F1-score) for the classification performance. All four indicator results showed that the MobileNetV3-Small model exhibited the best classification performance. In addition, the classification performance was in the order of EfficientNet-B0 > MobileNetV2 > NASNet-Mobile.

Tables 5 (A)–(D) present the confusion matrix, which presents the classification results of the four models. All four models showed accuracies of more than 90%. In the confusion matrix of each model, we could confirm that the classification for each skin group was performed well. Among the four models, MobileNetV3-Small most sensitively classified each group image.

There were several reasons for misclassification in the overall classification results. First, the aging skin had curved surfaces, and fine wrinkles were detected around them. When these fine wrinkles appeared densely in the segmentation image, the skin condition was perceived as younger. In other words, images of the old group were classified as those of the middle group (NASNet-Mobile, MobileNetV3-Small, and EfficientNet-B0). In addition, images of the middle group were classified as those of the young group (NASNet-Mobile, MobileNetV2, and EfficientNet-B0). Second, the wrinkles of young skin were shallow. Therefore, fine wrinkles could be disappeared by surrounding illumination or shadows. Hence, it was difficult to detect fine wrinkles in the image of young skin. The young skin structure appeared simple, consisting of only major wrinkles, such that images of the young group were classified as those of the middle group or the old group (NASNet-Mobile, MobileNetV2, and MobileNetV3-Small). In addition, the skin of the middle group had clearer wrinkles

TABLE 5. Confusion matrix of four CNN models for skin aging classification.

(A) CONFUSION MATRIX OF NASNET-MOBILE

		True labels		
		Young	**Middle**	**Old**
Predicted labels	**Young**	**39**	0	1
	Middle	2	**32**	6
	Old	0	1	**39**

(B) CONFUSION MATRIX OF MOBILENETV2

		True labels		
		Young	**Middle**	**Old**
Predicted labels	**Young**	**39**	1	0
	Middle	2	**32**	6
	Old	0	0	**40**

(C) CONFUSION MATRIX OF MOBILENETV3-SMALL

		True labels		
		Young	**Middle**	**Old**
Predicted labels	**Young**	**39**	1	0
	Middle	0	**35**	4
	Old	0	1	**39**

(D) CONFUSION MATRIX OF EFFICIENTNET-B0

		True labels		
		Young	**Middle**	**Old**
Predicted labels	**Young**	**40**	0	0
	Middle	2	**34**	4
	Old	1	1	**38**

than the young skin and had less prominent fine wrinkles caused by the skin curvature than the old skin. Therefore, in all classifiers, some images of the middle group were classified as those of the old group.

We classified the skin aging and condition using two-dimensional skin structure characteristics such as the pattern of wrinkles and shape of the microstructure. Therefore, misclassification was caused by curvature, shading, and noises that deform the skin microstructure. Nevertheless, all four CNN models exhibited good performance for the skin aging classification. There were characteristic differences in the skin microstructure according to aging. We can evaluate skin aging by analyzing the skin microstructure.

V. DISCUSSION

The microstructure of the skin surface has various elements such as wrinkles, wrinkled cells, texture, and anisotropy, which can help in the estimation of skin conditions. However, as time progresses, the microstructure of the skin shows both topological and morphological changes. Therefore, these aging characteristics make it difficult to analyze the microstructure of old skin. In addition, unlike dermoscopy images used in many previous skin-related studies, mobile images are unconstrained images exposed to the surrounding environmental noises. The mobile skin images showed significant topological changes on the skin surface related to aging. It is difficult to analyze the skin microstructure in the mobile images with the naked eye. Therefore, we captured the skin images of various aging types using a mobile device and then performed the skin microstructure segmentation and skin aging classification using CNN models.

First, we increased the accuracy of skin microstructure segmentation by adding a fusion module to the U-Net structure. Comparing the segmentation results of the watershed-algorithm-based method from a previous study and those of the deep learning models, FCN-8, UNet, and ResUNet++, and fusion UNet, we confirmed that the segmentation results of the fusion UNet model were most similar to the GT. The results of the watershed-algorithm-based method indicated under-segmentation as adjacent wrinkled cells were integrated by noises such as hair, and shading. FCN-8 estimated the wrinkles region wider than that of the GT and achieved the worst segmentation result. The results of UNet and ResUNet++ were similar to the GT, but had many noisy areas owing to over-segmentation around the wrinkles. In particular, ResUNet++ did not recognize the wrinkle structure when there was a significant shading. The results of the proposed fusion UNet were also similar to the GT and had less noise. In addition, the fusion UNet extracted fine wrinkles when the image was blurred or when the wrinkles were ambiguous owing to surrounding environmental noise such as illumination or shadows.

Next, we classified skin aging using segmented skin microstructure images. We used four mobile CNN models, NASNet-Mobile, MobileNetV2, MobileNetV3-Small, and EfficientNet-B0 for classification, and compared and analyzed their classification performance. All four models showed accuracies of more than 90%. MobileNetV3-Small showed the best classification performance with an accuracy of 94%. In the overall classified results, misclassification occurred when fine wrinkles appeared densely around the skin curvature in the old group images or when fine wrinkles were not precisely detected in the young group images.

Nevertheless, we confirmed differences in the skin microstructure characteristics with aging. Further, we evaluated and classified the skin aging or condition using these differences.

VI. CONCLUSION

Skin is an organ that reflects internal and external factors; therefore, observing and analyzing skin features such as wrinkles are important for evaluating current skin conditions and aging. In previous skin-related studies, special equipment such as a dermoscopy has been used to obtain skin images under limited conditions. It is possible to evaluate the morphological changes of the skin microstructure with aging by using these standardized skin images. However, it is difficult to observe other aging characteristics, such as topological changes. In contrast, mobile skin images include various environmental noises, but various types of skin aging can be observed, including morphological and topological changes in the broader skin area. Therefore, we performed skin analysis reflecting various types of skin aging by using mobile images. We used deep learning methods to detect various skin features that appear in old skin to young skin. We proposed a fusion U-Net model to perform skin microstructure segmentation and then classified microstructure-based skin aging using NASNet-Mobile, MobileNetV2, MobileNetV3-Small, and EffiecientNet-B0.

To more precisely extract the skin microstructure, we used the watershed-algorithm-based method, deep learning models (FCN-8, UNet, and ResUNet++), and proposed fusion UNet. The results of the fusion UNet were most similar to the GT. The fusion UNet also extracted fine wrinkles that were difficult to recognize with the naked eye. The remaining four segmentation methods resulted in various errors in skin images with surface curvature, hair, scars, illumination, and shading. The extracted skin microstructure had different skin characteristics related to aging. As a result of aging classification using these differences, MobileNetV3-Small showed a classification accuracy of 94%. Therefore, we confirmed that aging analysis is possible using the skin microstructure. In addition, the proposed method could be utilized in a mobile-based self-diagnosis system.

We evaluated skin aging evaluation using two-dimensional skin microstructure characteristics. However, the skin surface contains two-dimensional morphological features and three-dimensional features such as curvature and depth; therefore, an additional three-dimensional skin evaluation method is required for more accurate skin evaluations. In the future study, we will perform a three-dimensional skin analysis and increase the accuracy and precision of the skin evaluation.

REFERENCES

[1] J. M. Lagarde, C. Rouvrais, and D. Black, "Topography and anisotropy of the skin surface with ageing," *Skin Res. Technol.*, vol. 11, no. 2, pp. 110–119, May 2005.

[2] Y. Zou, E. Song, and R. Jin, "Age-dependent changes in skin surface assessed by a novel two-dimensional image analysis," *Skin Res. Technol.*, vol. 15, no. 4, pp. 399–406, 2009.

[3] M. A. Hamer, L. C. Jacobs, J. S. Lall, A. Wollstein, L. M. Hollestein, A. R. Rae, K. W. Gossage, A. Hofman, F. Liu, M. Kayser, T. Nijsten, and D. A. Gunn, "Validation of image analysis techniques to measure skin aging features from facial photographs," *Skin Res. Technol.*, vol. 21, no. 4, pp. 392–402, Nov. 2015.

[4] J. Rew, Y.-H. Choi, H. Kim, and E. Hwang, "Skin aging estimation scheme based on lifestyle and dermoscopy image analysis," *Appl. Sci.*, vol. 9, no. 6, p. 1228, Mar. 2019.

[5] J. Xie, L. Zhang, J. You, D. Zhang, and X. Qu, "A study of hand back skin texture patterns for personal identification and gender classification," *Sensors*, vol. 12, no. 7, pp. 8691–8709, Jun. 2012.

[6] Q. Gao, J. Yu, F. Wang, T. Ge, L. Hu, and Y. Liu, "Automatic measurement of skin textures of the dorsal hand in evaluating skin aging," *Skin Res. Technol.*, vol. 19, no. 2, pp. 145–151, May 2013.

[7] C. Trojahn, G. Dobos, M. Schario, L. Ludriksone, U. Blume-Peytavi, and J. Kottner, "Relation between skin micro-topography, roughness, and skin age," *Skin Res. Technol.*, vol. 21, no. 1, pp. 69–75, 2015.

[8] C. I. Moon and O. Lee, "Age-dependent skin texture analysis and evaluation using mobile camera image," *Skin Res. Technol.*, vol. 24, no. 3, pp. 490–498, 2018.

[9] G. Hong and O. Lee, "Three-dimensional reconstruction of skin disease using multi-view mobile images," *Skin Res. Technol.*, vol. 25, no. 4, pp. 434–439, 2019.

[10] C.-I. Moon and O. Lee, "Adaptive fine distortion correction method for stereo images of skin acquired with a mobile phone," *Sensors*, vol. 20, no. 16, p. 4492, Aug. 2020.

[11] J. Oh, H. Hong, Y. Cho, H. Yun, K. H. Seo, H. Kim, M. Kim, and O. Lee, "A reliable quasi-dense corresponding points for structure from motion," *KSII Trans. Internet Inf. Syst.*, vol. 14, no. 9, pp. 3782–3796, 2020.

[12] X. Yang, Z. Zeng, S. Yong Yeo, C. Tan, H. Liang Tey, and Y. Su, "A novel multi-task deep learning model for skin lesion segmentation and classification," 2017, *arXiv:1703.01025*.

[13] M. Saha and C. Chakraborty, "Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2189–2200, May 2018.

[14] Z. Wu, S. Zhao, Y. Peng, X. He, X. Zhao, K. Huang, X. Wu, W. Fan, F. Li, and M. Chen, "Studies on different CNN algorithms for face skin disease classification based on clinical images," *IEEE Access*, vol. 7, pp. 66505–66511, 2019.

[15] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," *IEEE Access*, vol. 8, pp. 7160–7172, 2020.

[16] A. Wagh, S. Jain, A. Mukherjee, E. Agu, P. C. Pedersen, D. Strong, B. Tulu, C. Lindsay, and Z. Liu, "Semantic segmentation of smartphone wound images: Comparative analysis of AHRF and CNN-based approaches," *IEEE Access*, vol. 8, pp. 181590–181604, 2020.

[17] M. A. Anjum, J. Amin, M. Sharif, H. U. Khan, M. S. A. Malik, and S. Kadry, "Deep semantic segmentation and multi-class skin lesion classification based on convolutional neural network," *IEEE Access*, vol. 8, pp. 129668–129678, 2020.

[18] A. Abu Mallouh, Z. Qawaqneh, and B. D. Barkana, "Utilizing CNNs and transfer learning of pre-trained models for age range classification from unconstrained face images," *Image Vis. Comput.*, vol. 88, pp. 41–51, Aug. 2019.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[20] T. L. B. Khanh, D.-P. Dao, N.-H. Ho, H.-J. Yang, E.-T. Baek, G. Lee, S.-H. Kim, and S. B. Yoo, "Enhancing U-Net with spatial-channel attention gate for abnormal tissue segmentation in medical imaging," *Appl. Sci.*, vol. 10, no. 17, p. 5729, Aug. 2020.

[21] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[22] D. Jha, P. H. Smedsrud, M. A. Riegler, D. Johansen, T. D. Lange, P. Halvorsen, and H. D. Johansen, "ResUNet++: An advanced architecture for medical image segmentation," in *Proc. IEEE Int. Symp. Multimedia (ISM)*, Dec. 2019, pp. 225–2255.

[23] T. J. Sorensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons," *Biol. Skar.*, vol. 5, pp. 1–34, 1948.
[24] P. Jaccard, "The distribution of the flora in the Alpine zone. 1," *New Phytologist*, vol. 11, no. 2, pp. 37–50, Feb. 1912.
[25] M.-P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. 12th Int. Conf. Pattern Recognit.*, vol. 1, Oct. 1994, pp. 566–568.
[26] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *J. Amer. Statist. Assoc.*, vol. 66, no. 336, pp. 846–850, 1971.
[27] B. Sathya and R. Manavalan, "Image segmentation by clustering methods: Performance analysis," *Int. J. Comput. Appl.*, vol. 29, no. 11, pp. 27–32, Sep. 2011.
[28] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8697–8710.
[29] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
[30] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324.
[31] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.

CHO-I MOON received the M.S. degree in biomedical engineering from the Department of Software Convergence, Soonchunhyang University, Asan-si, South Korea, in 2019, where she is currently pursuing the Ph.D. degree. Her research interests include the medical imaging processing, mobile applications, and computer vision.

ONSEOK LEE received the B.S. degree in electrical engineering and the M.S. and Ph.D. degrees in biomedical engineering from Korea University, Seoul, South Korea, in 2005, 2007, and 2011, respectively. He is currently an Assistant Professor with the Department of Medical IT Engineering, Soonchunhyang University, Asan-si, South Korea. His research interests include the medical imaging analysis and computer vision.