Automated Facial Wrinkle Segmentation Scheme Using UNet++

Hyeonwoo Kim[1], Junsuk Lee[1], Jehyeok Rew[2], and Eenjun Hwang[1*]
[1] School of Electrical Engineering, Korea University
Seoul, Republic of Korea
[e-mail: guihon12@korea.ac.kr, lilemoest48@korea.ac.kr, ehwang04@korea.ac.kr]
[2] Department of Data Science, Duksung Women's University
Seoul, Republic of Korea
[e-mail: jhrew@duksung.ac.kr]
[*]Corresponding author: Eenjun Hwang

Abstract

Facial wrinkles are widely used to evaluate skin condition or aging for various fields such as skin diagnosis, plastic surgery consultations, and cosmetic recommendations. In order to effectively process facial wrinkles in facial image analysis, accurate wrinkle segmentation is required to identify wrinkled regions. Existing deep learning-based methods have difficulty segmenting fine wrinkles due to insufficient wrinkle data and the imbalance between wrinkle and non-wrinkle data. Therefore, in this paper, we propose a new facial wrinkle segmentation method based on a UNet++ model. Specifically, we construct a new facial wrinkle dataset by manually annotating fine wrinkles across the entire face. We then extract only the skin region from the facial image using a facial landmark point extractor. Lastly, we train the UNet++ model using both dice loss and focal loss to alleviate the class imbalance problem. To validate the effectiveness of the proposed method, we conduct comprehensive experiments using our facial wrinkle dataset. The experimental results showed that the proposed method was superior to the latest wrinkle segmentation method by 9.77%p and 10.04%p in IoU and F1 score, respectively.

1. Introduction

Facial wrinkles appear in the form of fine lines or creases on the skin of the human face [1]. These wrinkles are a natural sign of aging, but they are also affected by many other factors, including sun exposure, smoking, and changes in weight [2,3]. Wrinkles are one of the key targets in facial analysis because they are used as the basis for evaluating skin aging or offering skin and beauty-related recommendations [4]. Wrinkles are a particularly useful visual indicator of aging because they reflect changes in skin elasticity, collagen levels, and overall tissue integrity. Therefore, accurate analysis of facial wrinkles can be used as the foundation for the objective and effective assessment of the efficacy of various cosmetic treatments and skincare interventions.

In facial image analysis, a variety of facial wrinkle segmentation methods have been proposed [5-7], which can be broadly divided into traditional image processing and deep learning-based approaches. Traditional image processing-based methods mostly rely on handcrafted features and rule-based algorithms [8-10]. While these methods have demonstrated effective wrinkle segmentation performance for preprocessed facial images, they have limitations in detecting complex and diverse wrinkle patterns. On the other hand, deep learning-based methods demonstrated remarkable segmentation performance across various image processing domains because they can learn intricate features from data [11-13]. Consequently, there has been a recent surge in studies employing these approaches for accurate wrinkle segmentation [14-17]. Nevertheless, it remains difficult to construct effective deep learning-based wrinkle segmentation models due to the lack of facial wrinkle data and the class imbalance between wrinkles and other facial features. In particular, the wrinkle class imbalance often causes models to focus on areas other than wrinkles, which account for most of a face image, resulting in inaccurate predictions due to overfitting.

In this paper, we propose a novel facial wrinkle segmentation scheme to overcome these problems. In particular, we construct a new facial wrinkle dataset by manually annotating fine wrinkles in facial images from a known face dataset. We then extract regions of interest where wrinkles appear in the facial images using a facial landmark point extractor. Finally, we train a UNet++ model [18] using both dice loss [19] and focal loss [20] to alleviate the imbalance between the wrinkles and non-wrinkle elements. Dice loss and focal loss punish false predictions and simultaneously down-weigh non-wrinkle elements while concentrating on the wrinkles. To demonstrate the effectiveness of the proposed scheme, we perform comparative experiments using our wrinkle dataset.

The structure of this paper is as follows. Section 2 provides a comprehensive overview of previous research on wrinkle segmentation. In Section 3, our proposed method is described, while Section 4 summarizes its performance in a series of comparative experiments. Finally, Section 5 presents the conclusions of our paper.

2. Related Works

Facial wrinkle segmentation is a pixel-level classification task for detailed representation of facial wrinkles. It involves identifying and isolating regions in facial images that show fine lines, creases, and textural irregularities. In this section, we summarize previous studies on facial wrinkle segmentation by categorizing them into image processing-based and deep learning-based approaches. We examine their strengths and limitations and briefly outline how our proposed method overcomes these limitations.

2.1 Image Processing-based Wrinkle Segmentation

Image processing-based facial wrinkle segmentation methods detect wrinkle patterns in facial images using predefined image filters. For instance, Batool et al. [8] used Gabor filter banks [21] to highlight discontinuities in skin texture and exploited the image morphology to localize wrinkle shapes. In addition, Ng et al. [9] proposed a hybrid Hessian filter (HHF) based on directional gradients and a ridge-valley Gaussian kernel and developed Hessian line tracking (HLT) [5] for facial wrinkle segmentation using HHF. Similarly, Yap et al. [10] created a facial wrinkle annotator based on HHF and showed that it can segment both fine and coarse wrinkles. However, while these image processing-based methods can effectively segment wrinkles, it is difficult to optimize their parameters due to the wide range of image capturing conditions, including differences in lighting and camera angles.

2.2 Deep Learning-based Facial Wrinkle Segmentation

Recently, deep learning-based methods have been proposed that are robust to changes in image conditions by directly learning the wrinkle patterns of facial images. Li et al. [15] proposed a nasolabial fold segmentation method that combines an object detection network with a semantic segmentation network. The nasolabial folds are the two skin folds that run from each side of the nose to the corners of the mouth. They adopted faster region-based convolutional neural networks (Faster R-CNNs) [22] to detect the nasolabial region in whole face images. They then used a global convolution network (GCN) [23] for nasolabial fold segmentation. Similarly, Umirzakova and Whangbo [14] proposed a method for nasolabial folds segmentation in whole face images using a nested CNN. They incorporated attention blocks within skip connections at multiple scales to address the imbalanced spatial size of the intermediate feature maps.

Other studies have looked at segmenting wrinkles around the forehead and eyes. For example, Kim et al. [16] introduced a new training strategy for segmenting wrinkles in facial images around the forehead and eyes. To generate wrinkle label data, they augmented a roughly labeled wrinkle data by multiplying a texture map extracted using a Gaussian filter and adaptive thresholding [24]. They then trained a U-Net model [25] to segment wrinkles in around the forehead and eyes using the augmented wrinkle labels. As a follow-up study, they also proposed a method to simultaneously detect areas of wrinkles and pores using this learning strategy [17]. These deep learning-based wrinkle segmentation methods have demonstrated good performance and have been employed in various research scenarios. For example, these methods have been used to extract wrinkle regions from facial images for seamless wrinkle removal through inpainting [26], and to detect wrinkles in the forehead area to assess the need for filler injections [7].

However, existing deep learning-based wrinkle segmentation methods are limited to specific facial regions, such as the forehead, eyes, and nasolabial folds, due to the challenges and high costs associated with creating facial wrinkle datasets. Moreover, expanding the target area for wrinkle segmentation significantly increases the ratio of wrinkles to non-wrinkles, leading to a data imbalance that degrades model performance. To mitigate these issues, we manually annotated ground truth wrinkles on whole face images and isolated the skin regions from the face images to minimize the non-wrinkle elements. We also used dice loss and focal loss to train our model to address the data imbalance problem.
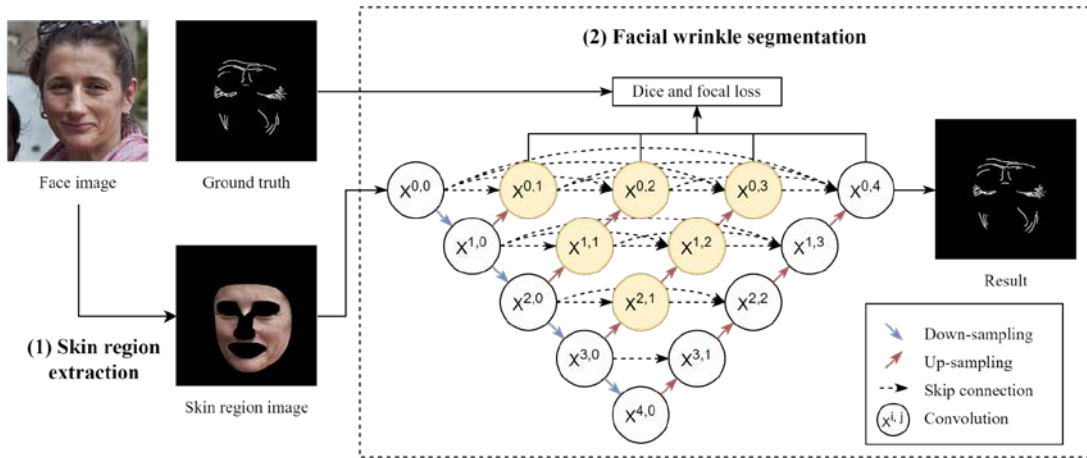
Fig. 1. Overview of the proposed facial wrinkle segmentation method.

3. Methodology

Fig. 1 presents an overview of the proposed facial wrinkle segmentation method. It consists of two steps: (1) extracting skin regions from face images, and (2) segmenting facial wrinkles using a UNet++ model with dice loss and focal loss. Before presenting a detailed description of each component, we first briefly describe the facial wrinkle dataset we constructed for model training and experiments.

Fig. 2. Samples of our wrinkle dataset.

3.1 Construction of the Facial Wrinkle Dataset

Because there are no public facial wrinkle datasets containing fine wrinkles, we created a new facial wrinkle dataset by manually annotating wrinkles on facial images using the public facial image dataset Flickr-Faces-HQ (FFHQ) [27] and used it to train our wrinkle segmentation model. FFHQ contains 70,000 high-quality face images. From the dataset, we selected 1,000 face images covering a diverse range of ages and carefully annotated the facial wrinkles, to capture various shapes, sizes, and complexities. **Fig. 2** shows examples of original face images and their wrinkle annotations. The wrinkle annotation results were used as the ground truth in model training and evaluation. Our wrinkle dataset is available at the website https://github.com/jun01pd2015/wrinkle_dataset.

3.2 Skin Region Extraction

In this step, we extracted only the skin regions from the facial images to focus on the wrinkle segmentation of facial skin. To represent only the regions where wrinkles may appear, we isolated the skin area from each face image by the removing landmarks and background from the image. **Fig. 3** presents the steps for skin region extraction. A face mesh consisting of 468 facial landmark points was first constructed from the face image using a face landmark extractor [28,29]. A facial skin mask representing only the skin area was created by removing areas containing the eyes, nose, mouth, and background from the face image based on the extracted facial landmark points. Finally, a facial skin image was constructed from the original face image by multiplying the original face image with the skin mask. By using only skin regions as input, our facial wrinkle segmentation model focused more on learning wrinkle patterns without the noise generated by non-wrinkle elements.
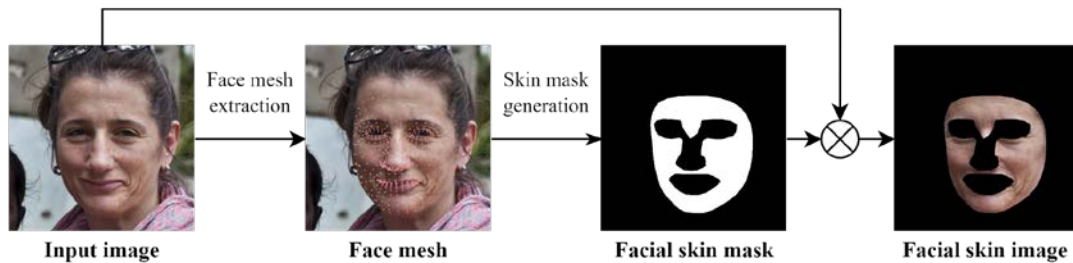

Fig. 3. Skin region extraction process.

3.3 Facial Wrinkle Segmentation Model

3.3.1 Model Architecture

The proposed model for facial wrinkle segmentation is based on UNet++ [18]. UNet++ is an extension of the U-Net [25] architecture commonly employed in medical image segmentation tasks. U-Net consists of an encoder and a decoder. The encoder extracts low-level features from input images, and the decoder utilizes these features to generate the desired output. U-Net incorporates skip connections to transmit the output from each encoder layer to the corresponding decoder layer to minimize information loss. UNet++ greatly expands these skip connections by adding iterative convolution blocks to reduce the semantic differences between the feature maps of the encoder subnetwork and the decoder subnetwork. UNet++ also

introduces deep supervision [30] to address the gradient vanishing problem by enabling direct weight updates based on the gradient of the intermediate layer. Since deep supervision considers the outputs of both the last layer and intermediate layers in the loss calculation, it integrates feature maps at multiple resolutions, improving overall network stability and accuracy of capturing intricate patterns and multi-scale structures. For the UNet++ encoder, we adopt ResNeXt-50 [31] as the backbone network. In the wrinkle segmentation, the encoder down-samples skin region images to extract wrinkle feature maps with contextual information. The down-sampled wrinkle feature maps are then propagated to the corresponding layer of sub-networks through densely nested skip pathways. Finally, the decoder up-samples feature maps from the corresponding layer of sub-networks to generate wrinkle masks.

3.3.2 Loss Function

Facial wrinkles usually occur in specific regions, such as around the eyes or mouth. As a result, wrinkle regions account for a significantly lower portion of the entire face compared to the region without wrinkles. This data imbalance can bias the model towards the higher-proportion classes during training, leading to a poor performance. The use of cross-entropy loss, which computes classification loss at the pixel level, renders the model particularly susceptible to this type of data imbalance. To solve this problem, we use two loss functions: dice loss [19] and focal loss [20]. Dice loss penalizes incorrect predictions, while focal loss down-weights non-wrinkle elements and focuses on wrinkle areas. Dice loss is a region-based loss that quantifies the similarity between the predicted segmentation mask and the ground truth mask, highlighting the overlapping regions. This loss, which is defined using (1), effectively alleviates the class imbalance by punishing incorrect predictions.

$$L_D(p, q) = 1 - \frac{2 \times \sum_{i=1}^{H \times W}(p_i \times q_i)}{\sum_{i=1}^{H \times W} p_i^2 + \sum_{i=1}^{H \times W} q_i^2}, \tag{1}$$

where p and q denote the predicted mask and ground truth mask, respectively, while H and W represent the height and width of the input image, respectively.

Focal loss is a variation of standard cross entropy loss that is designed to reduce the relative loss for well-classified examples and increase the relative loss for difficult or misclassified examples. It enhances minority class instance detection and mitigates learning bias towards the majority class by concentrating on challenging examples. In facial wrinkle segmentation, non-wrinkle elements are easy to distinguish (i.e., easy negatives), but wrinkles are difficult to identify (i.e., hard positives). Therefore, focal loss assigns low weights to non-wrinkle elements and high weights to wrinkle elements during training. Focal loss is defined as shown in (2):

$$L_F(prob) = \begin{cases} -\sum_{i=1}^{H \times W} \alpha(1 - prob_i)^\gamma \log(prob_i), & y = 1 \ (wrinkle) \\ -\sum_{i=1}^{H \times W} (1 - \alpha) prob_i{}^\gamma \log(1 - prob_i), & y = 0 \ (non\_wrinkle) \end{cases}, \tag{2}$$

where α is the weighting factor, γ is the focusing parameter, and $prob$ represents the predicted probability of the correct class. In the present study, we set the weighting factor to 0.75 and the focusing parameter to 2, following the environmental setting in [20].

4. Experiments

4.1 Environmental Setup

To evaluate the performance of the proposed scheme, we performed various experiments using our wrinkle dataset. In the experiments, we split the dataset into training and test datasets at a ratio of 9:1. Our segmentation method was implemented with the Pytorch framework [32]. We implemented the UNet++ [18] architecture with ResNeXt-50 [31] and trained it over 500 epochs using the Adam [33] optimizer with a learning rate of 0.0001. During the training process, the input image was resized to 512×512 and augmented with random horizontal flipping and random rotating from –45° to 45° to prevent overfitting. In addition, we set the threshold probability of the segmentation map to 0.5. All experiments were conducted on a Nvidia Titan RTX gpu. We reported the highest performance observed within an acceptable training period for each experiment.

4.2 Performance Evaluation

In this section, we qualitatively and quantitatively compared the proposed method with existing image processing-based methods [8,10] and a deep learning method [26]. First, **Fig. 4** shows their qualitative comparison results. The proposed model performed much better than other compared methods in terms of wrinkle segmentation, accurately reproducing the wrinkles of the ground truth while suppressing the reproduction of non-wrinkle elements. This indicates that our model is capable of more accurately and robustly learning wrinkle patterns. In contrast, the traditional deep learning method often segmented non-wrinkle elements as wrinkles, and the image processing-based methods exhibited a poor performance in segmenting all facial wrinkles. In particular, they detected both non-wrinkle elements and thick edges or ridges as wrinkles.
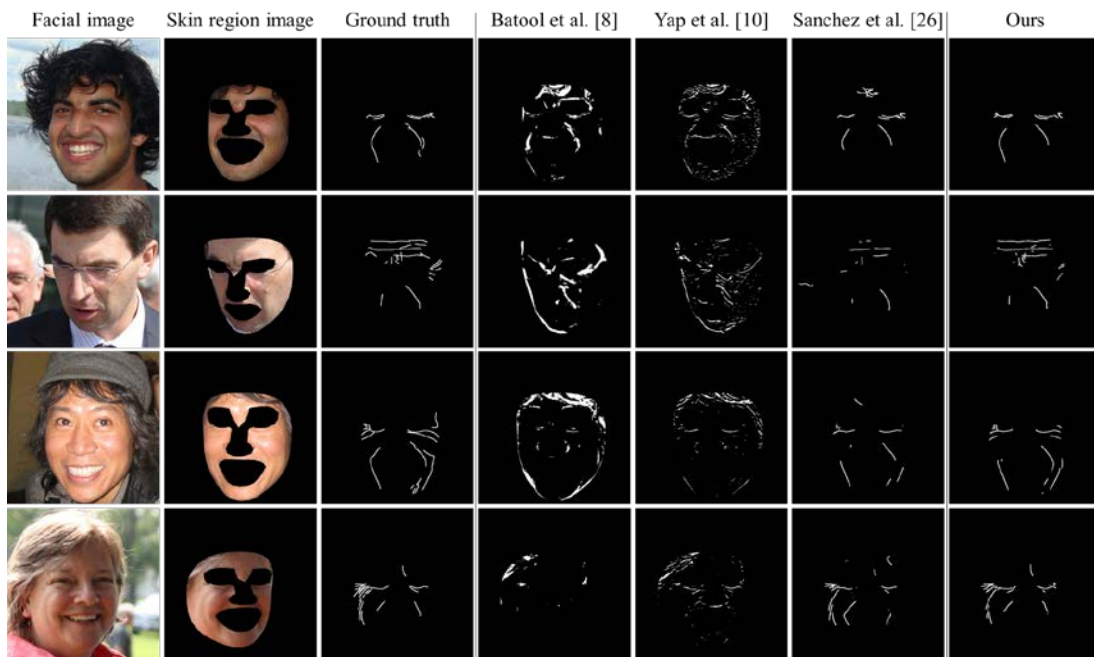

Fig. 4. Qualitative comparisons of facial wrinkle segmentation methods.

Next, we use a variety of metrics, including pixel accuracy, intersection over union (IoU), sensitivity, precision, F1 score, and specificity to quantitatively compare the efficacy and accuracy of wrinkle segmentation methods. Pixel accuracy measures the number of pixels correctly classified as wrinkles or non-wrinkles as a proportion of the total number of pixels. It can be defined using (3):

$$Pixel\ accuracy\ = \frac{TP + TN}{TP + TN + FP + FN},\quad\quad (3)$$

where TP and TN are the number of true positives (correctly identified wrinkle pixels) and true negatives (non-wrinkle pixels correctly identified), respectively, and FP and FN are the number of false positives (non-wrinkle pixels incorrectly classified as wrinkle) and false negatives (wrinkle pixels incorrectly classified as non-wrinkle), respectively. IoU is widely used in segmentation to evaluate the performance by measuring the overlap between the predicted area and the actual area (i.e., the ground truth). IoU is calculated using (4):

$$IoU\ = \frac{TP}{TP + FP + FN}\quad\quad (4)$$

Eq. (4) measures the relationship between correctly predicted object pixels and the total set of actual and predicted pixels. The IoU ranges from 0 and 1, where 1 indicates a perfect overlap between the predicted and actual areas, and 0 indicates no overlap. Sensitivity, defined by (5), represents the proportion of actual wrinkle pixels that are correctly identified as winkles, highlighting the model's ability to detect target pixels. The metric is defined as

$$Sensitivity\ = \frac{TP}{TP + FN}\quad\quad (5)$$

Precision accesses the proportion of predicted wrinkle pixels that are correctly classified, reflecting the accuracy of a model in identifying true wrinkle regions without over-segmentation. The metric is defined as shown in (6):

$$Precision\ = \frac{TP}{TP + FP}\quad\quad (6)$$

The F1 score synthesizes the balance between sensitivity and precision, emphasizing the accuracy of correctly identified wrinkle pixels. The metric is defined using (7):

$$F1\ score\ = \frac{2 \times Sensitivity \times Precision}{Sensitivity + Precision} = \frac{2TP}{2TP + FP + FN}\quad\quad (7)$$

Specificity calculates the proportion of correctly identified non-wrinkle pixels, indicating a model's capability to accurately exclude non-relevant areas. This metric is defined as

$$Specificity\ = \frac{TN}{TN + FP}\quad\quad (8)$$

Table 1. Quantitative comparison of wrinkle segmentation methods.

Segmentation methods	Pixel accuracy	IoU	Sensitivity	Precision	F1 score	Specificity
Batool et al. [8]	0.9623	0.0818	0.2628	0.1262	0.1479	0.9728
Yap et al. [10]	0.979	0.101	0.1737	0.2275	0.181	0.9909
Sanchez et al. [26]	0.9889	0.3517	0.4289	**0.6726**	0.5153	**0.997**
Ours	**0.9896**	**0.4494**	**0.6019**	0.6435	**0.6157**	0.9952

Table 1 summarizes the results comparing our proposed method with other wrinkle segmentation methods from Batool et al. [8], Yap et al. [10], and Sanchez et al. [26]. Our method achieved the highest pixel accuracy (0.9896), indicating that it correctly classified pixels as wrinkles or non-wrinkles. Sanchez et al.'s model [26] followed closely behind with 0.9889, while Batool et al. [8] and Yap et al. [10] exhibited a lower accuracy. Our method was also the best-performing approach in terms of IoU with 0.4494, significantly outperforming the other methods. This suggests that our method was superior in accurately delineating the wrinkle regions. The low IoU of Batool et al. [8] reflected a strong discrepancy between the predicted and actual wrinkle regions. In terms of sensitivity and precision, our method demonstrated a balanced performance (0.6019 and 0.6435, respectively), representing the effective identification of wrinkle pixels while maintaining a lower rate of false positives. Sanchez et al. [26] achieved a higher precision (0.6726), but our method achieved a better balance between detecting most wrinkles (i.e., sensitivity) and accurately classifying wrinkle pixels (i.e., precision). Our method also achieved the highest F1 score (0.6157), which balanced sensitivity and precision, representing a superior overall performance in terms of wrinkle detection accuracy. Finally, while our method demonstrated a slightly lower specificity (0.9952) compared to Sanchez et al. [26] (0.997) and Yap et al. [10] (0.9909), it remained high, indicating a strong ability to correctly identify non-wrinkle regions.

Table 2. Quantitative comparison of loss function.

Loss functions	Pixel accuracy	IoU	Sensitivity	Precision	F1 score	Specificity
BCE	0.9894	0.3363	0.3801	0.7602	0.4966	0.9981
Dice	0.9902	0.4358	0.5366	0.7003	0.6021	0.9968
Focal	0.9899	0.4148	0.5041	0.7025	0.5816	0.997
BCE + Focal	0.9901	0.3684	0.4091	**0.7868**	0.5309	**0.9984**
BCE + Dice	**0.9902**	0.4436	0.5669	0.6845	0.6151	0.9963
Dice + Focal	0.9896	**0.4494**	**0.6019**	0.6435	**0.6157**	0.9952
BCE + Dice + Focal	0.9899	0.4386	0.5584	0.6725	0.6044	0.9962

4.3 Comparative Analysis of Loss Functions

This section summarizes the impact of the loss function on model performance by training the model with various combinations of binary cross-entropy (BCE), dice loss, and focal loss. **Table 2** presents a summary of the results for multiple metrics for each loss function combination. The *Dice + Focal* combination achieved the highest sensitivity (0.6019), demonstrating its superior ability to accurately identify wrinkle regions. This high sensitivity is particularly crucial in datasets with a class imbalance because failing to detect the minority class (i.e., wrinkles in the present study) can significantly affect a model's utility. Moreover,

with an IoU of 0.4494, the *Dice + Focal* combination outperformed the other configurations. IoU directly measures the accuracy in the context of a spatial imbalance by evaluating the overlap between predicted and actual segmentation areas, making it an important metric for segmentation tasks. Although pixel accuracy and specificity were slightly lower with the *Dice + Focal* combination than for the other loss function combinations, they remained high. On the other hand, the *Dice + Focal* combination was considerably lower for precision than the *BCE + Focal* combination, which reached 0.7868. However, the *Dice + Focal* combination had the highest F1 score (0.6157), demonstrating its effectiveness in maintaining a balance between correctly identifying wrinkle regions and minimizing over-segmentation. This analysis highlights the importance of choosing the appropriate loss function combination for handling specific challenges, such as the class imbalance in wrinkle segmentation tasks. Its outstanding performance in terms of the F1 score, sensitivity, and IoU suggests that the combination of dice loss and focal loss was the most suitable for training wrinkle segmentation models.

5. Conclusion

In this paper, we proposed an automated facial wrinkle segmentation scheme that can effectively segment wrinkles over the entire face. To achieve this, we constructed a new facial wrinkle dataset with deep and shallow wrinkles for various age groups. We also extracted skin regions from facial images to focus on the wrinkle patterns of facial skin. To overcome the class imbalance between wrinkles and non-wrinkle elements, we used both dice loss and focal loss to train the UNet++ model. Comparative experiments using our wrinkle dataset demonstrated that our proposed method significantly outperformed existing facial wrinkle segmentation methods. In future work, we plan to expand our research into facial skin analysis. This will include identifying wrinkle patterns across various demographic features, including age, race, and gender, using the proposed method.

References

[1]   P. K. Mukherjee, N. Maity, N. K. Nema, and B. K. Sarkar, "Bioactive compounds from natural resources against skin aging," *Phytomedicine*, vol.19, no.1, pp.64-73, 2011. Article(CrossRefLink)

[2]   P. Quatresooz, L. Thirion, C. Piérard-Franchimont, and G.E. Piérard, "The riddle of genuine skin microrelief and wrinkles," *International Journal of Cosmetic Science*, vol.28, no.6, pp.389-395, 2006. Article(CrossRefLink)

[3]   O. F. Osman, R. M. I. Elbashir, I. E. Abbass, C. Kendrick, M. Goyal, and M. H. Yap, "Automated assessment of facial wrinkling: A case study on the effect of smoking," in *Proc. of 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp.1081-1086, 2017. Article(CrossRefLink)

[4]   M. H. Yap, N. Batool, C. C. Ng, M. Rogers, and K. Walker, "A Survey on Facial Wrinkles Detection and Inpainting: Datasets, Methods, and Challenges," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol.5, no.4, pp.505-519, 2021. Article(CrossRefLink)

[5]   G. O. Cula, P. R. Bargo, A. Nkengne, and N. Kollias, "Assessing facial wrinkles: automatic detection and quantification," *Skin Research and Technology*, vol.19, no.1, pp.e243-e251, 2013. Article(CrossRefLink)

[6]   C. C. Ng, M. H. Yap, N. Costen, and B. Li, "Wrinkle Detection Using Hessian Line Tracking," *IEEE Access*, vol.3, pp.1079-1088, 2015. Article(CrossRefLink)

[7] A. Alrabiah, M. Alduailij, and M. Crane, "Computer-based Approach to Detect Wrinkles and Suggest Facial Fillers," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol.10, no.9, pp.319-325, 2019. Article(CrossRefLink)

[8] N. Batool and R. Chellappa, "Fast detection of facial wrinkles based on Gabor features using image morphology and geometric constraints," *Pattern Recognition*, vol.48, no.3, pp.642-658, 2015. Article(CrossRefLink)

[9] C. C. Ng, M. H. Yap, N. Costen, and B. Li, "Automatic Wrinkle Detection Using Hybrid Hessian Filter," in *Proc. of Computer Vision -- ACCV 2014, 12th Asian Conference on Computer Vision*, vol.9005, pp.609-622, 2015. Article(CrossRefLink)

[10] M. H. Yap, J. Alarifi, C. C. Ng, N. Batool, and K. Walker, "Automated Facial Wrinkles Annotator," in *Proc. of Computer Vision – ECCV 2018 Workshops*, vol.11132, pp.676-680, 2019. Article(CrossRefLink)

[11] H. Kim, C. Kim, H. Kim, S. Cho, and E. Hwang, "Panoptic blind image inpainting," *ISA Transactions*, vol.132, pp.208-221, 2023. Article(CrossRefLink)

[12] H. Kim, H. Kim, J. Shim, and E. Hwang, "A robust kinship verification scheme using face age transformation," *Computer Vision and Image Understanding*, vol.231, 2023. Article(CrossRefLink)

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol.60, no.6, pp.84-90, 2017. Article(CrossRefLink)

[14] H. Li, K. Wu, H. Cheng, C. Gu, and X. Guan, "Nasolabial Folds Extraction based on Neural Network for the Quantitative Analysis of Facial Paralysis," in *Proc. of 2018 2nd International Conference on Imaging, Signal Processing and Communication (ICISPC)*, pp.54-58, 2018. Article(CrossRefLink)

[15] S. Umirzakova and and T. K. Whangbo, "Nasolabial Wrinkle Segmentation Based on Nested Convolutional Neural Network," in *Proc. of 2021 International Conference on Information and Communication Technology Convergence (ICTC)*, pp.483-485, 2021. Article(CrossRefLink)

[16] S. Kim, H. Yoon, J. Lee, and S. Yoo, "Semi-automatic Labeling and Training Strategy for Deep Learning-based Facial Wrinkle Detection," in *Proc. of 2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)*, pp.383-388, 2022. Article(CrossRefLink)

[17] H. Yoon, S. Kim, J. Lee, and S. Yoo, "Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis," *Diagnostics*, vol.13, no.11, 2023. Article(CrossRefLink)

[18] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," in *Proc. of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, DLMIA ML-CDS 2018*, vol.11045, pp.3-11, 2018. Article(CrossRefLink)

[19] W.R. Crum, O. Camara, and D.L.G. Hill, "Generalized Overlap Measures for Evaluation and Validation in Medical Image Analysis," *IEEE Transactions on Medical Imaging*, vol.25, no.11, pp.1451-1461, 2006. Article(CrossRefLink)

[20] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.42, no.2, pp.318-327, 2020. Article(CrossRefLink)

[21] D. Gabor, "Theory of communication. Part 1: The analysis of information," *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol.93, no.26, pp.429-441, 1946. Article(CrossRefLink)

[22] R. Girshick, "Fast R-CNN," in *Proc. of 2015 IEEE International Conference on Computer Vision (ICCV)*, pp.1440-1448, 2015. Article(CrossRefLink)

[23] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large Kernel Matters — Improve Semantic Segmentation by Global Convolutional Network," in *Proc. of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1743-1751, 2017. Article(CrossRefLink)

[24] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image," *Journal of Graphics Tools*, vol.12, no.2, pp.13-21, 2007. Article(CrossRefLink)

[25] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional Networks for Biomedical Image Segmentation," in *Proc. of Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol.9351, pp.234-241, 2015. Article(CrossRefLink)

[26] M. Sanchez, G. Triginer, C. Ballester, L. Raad, and E. Ramon, "Photorealistic Facial Wrinkles Removal," in *Proc. of 16th Asian Conference on Computer Vision – ACCV 2022 Workshops*, vol.13848, pp.117-133, 2023. Article(CrossRefLink)

[27] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.43, no.12, pp.4217-4228, 2021. Article(CrossRefLink)

[28] Y. Kartynnik, A. Ablavatski, I. Grishchenko, and M. Grundmann, "Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs," *arXiv:1907.06724*, 2019. Article(CrossRefLink)

[29] I. Grishchenko, A. Ablavatski, Y. Kartynnik, K. Raveendran, and M. Grundmann, "Attention Mesh: High-fidelity Face Mesh Prediction in Real-time," *arXiv:2006.10962*, 2020. Article(CrossRefLink)

[30] C. Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-Supervised Nets," in *Proc. of the Eighteenth International Conference on Artificial Intelligence and Statistics, PMLR*, vol.38, pp.562-570, 2015. Article(CrossRefLink)

[31] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," in *Proc. of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.5987-5995, 2017. Article(CrossRefLink)

[32] P. Iakubovskii, "Segmentation models PyTorch," *GitHub repository*, 2019. Article(CrossRefLink)

[33] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proc. of the 3rd International Conference for Learning Representations, San Diego*, 2015. Article(CrossRefLink)

Hyeonwoo Kim received his B.S. degree in Information and Communication Engineering from Hansung University, Seoul, Korea, in 2017. Currently, he is a doctoral student at the School of Electrical Engineering, Korea University, Seoul, Korea. His research interests include Face image analysis, deep learning, computer vision, and generative models.

Junsuk Lee received his B.S. degree in School of Health and Environmental Science and Interdisciplinary Major in Artificial Intelligence from Korea University, Seoul, Korea, in 2022. He received his M.S. degree in School of Electrical Engineering from Korea University, Seoul, Korea, in 2024. His research interests include deep learning, computer vision, image inpainting, and image segmentation.

Jehyeok Rew received his B.S. degree in Electronics and Radio Engineering from KyungHee University, Suwon, Korea, in 2012, and his integrated Ph.D. degree in School of Electrical Engineering from Korea University, Seoul, Korea, in 2020. Currently, he is a member of the faculty at the Department of Data Science, Duksung Women's University, Seoul, South Korea. His current research interests include GIS, information retrieval, big data processing, and machine learning.

Eenjun Hwang received his B.S. and M.S. degrees in Computer Engineering from Seoul National University, Seoul, Korea, in 1988 and 1990, respectively; and his Ph.D. degree in Computer Science from the University of Maryland, College Park, in 1998. Currently, he is a member of the faculty at the School of Electrical Engineering, Korea University, Seoul, South Korea. His current research interests include database, multimedia systems, information retrieval, big data processing, and machine learning.