A DEEP LEARNING APPROACH FOR ACCURATE FACIAL WRINKLE SEGMENTATION USING UNET++ MODEL WITH DICE AND FOCAL LOSS FUNCTIONS

Bui Nguyen Tuan Anh*, Bui Ngoc Tien, Nguyen Minh Trung, Hoang Minh Hau

Dong Nai Technology University.

Corresponding author: *Bui Nguyen Tuan Anh, email: anh.buinguyentuan@dntu.edu.vn*

GENERAL INFORMATION

Received date: 23/03/2024

Revised date: 07/05/2024

Accepted date: 11/07/2024

KEYWORD

Deep learning;

Dice loss;

Facial wrinkle segmentation;

Focal loss;

Unet++.

ABSTRACT

Facial wrinkle segmentation is essential for various applications in dermatology and cosmetology, yet existing methods often struggle with accurate delineation due to limited dataset diversity and class imbalance. In this paper, we propose a novel facial wrinkle segmentation method based on the Unet++ model, enhanced with dice and focal loss functions. Our approach begins with the construction of an enriched wrinkle dataset sourced from the Flickr-Faces-HQ dataset, ensuring diversity in wrinkle types and complexities. We then introduce a skin region extraction technique to isolate relevant facial areas, enhancing segmentation accuracy. The Unet++ model is employed for wrinkle segmentation, leveraging its encoder-decoder architecture and densely nested skip pathways to capture fine wrinkle details. By integrating dice and focal loss functions, our method effectively addresses class imbalance and improves segmentation performance. Experimental results demonstrate the superiority of our approach in both qualitative and quantitative evaluations, showcasing enhanced wrinkle extraction capabilities and superior segmentation accuracy compared to existing methods. Overall, our study advances the field of facial wrinkle segmentation, offering a robust and reliable method for accurate wrinkle delineation in facial images.

1. INTRODUCTION

Facial wrinkle segmentation, a crucial task in image analysis and dermatology, involves delineating fine lines and wrinkles within facial images. Accurate segmentation of these features facilitates various applications such as age estimation, cosmetic surgery planning, and skin condition assessment (M. Kim & Lee, 2021; Zheng et al., 2022). While traditional methods have been used for wrinkle segmentation, recent advancements in deep learning offer promising avenues for more precise and automated segmentation techniques.

However, challenges persist in achieving accurate wrinkle segmentation due to limited and unbalanced datasets. Existing datasets often lack diversity in wrinkle types and are skewed towards non-wrinkle regions, leading to suboptimal performance of segmentation models. In this context, our paper proposes a novel method to address these challenges and enhance the accuracy of facial wrinkle segmentation (Lboukili et al., 2023; Nader et al., 2022).

Our approach comprises several key components aimed at improving dataset quality, focusing on relevant facial regions, and leveraging advanced deep learning techniques. Specifically, we introduce a method for enhancing wrinkle datasets by augmenting existing data with diverse wrinkle types and complexities (Mukasheva et al., 2023). Additionally, we develop a technique for extracting skin regions from facial images, allowing segmentation models to focus exclusively on wrinkle patterns while eliminating noise from non-relevant areas (Gowroju et al., 2022; Yoon et al., 2023).

Central to our method is the adoption of the Unet++ architecture, a state-of-the-art deep learning model for semantic segmentation tasks. We enhance the training process by incorporating both dice loss and focal loss functions, which effectively address the imbalance between wrinkle and non-wrinkle data (Dhalla et al., 2023; S. Kim et al., 2023; Moon & Lee, 2022). Through comprehensive experimentation on our enhanced dataset, we demonstrate the superiority of our method in accurately segmenting facial wrinkles across diverse images. We provide detailed insights into each component of our proposed method, present experimental results, and discuss implications and future directions. Our study contributes to advancing the field of facial wrinkle segmentation and lays the groundwork for more effective and robust segmentation techniques in the future (Yang et al., 2024).

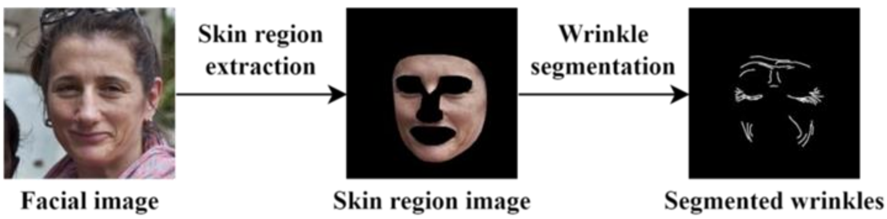

Figure. 1. illustrates an overview of the proposed facial wrinkle segmentation methodology.

2. METHODOLOGY

Our methodology for facial wrinkle segmentation is structured into several interconnected stages, each meticulously designed to address specific challenges and enhance the accuracy of the segmentation process. The first stage involves the construction of an enhanced wrinkle dataset. To overcome the limitations of existing datasets, we annotate wrinkles on images sourced from the Flickr-Faces-HQ (FFHQ) dataset. This meticulous annotation process ensures the inclusion of diverse wrinkle types, sizes, and complexities, thereby enhancing the quality and representativeness of the training data.

Following dataset construction, we proceed to the skin region extraction stage. Here, we develop a method to isolate skin regions within facial images, removing extraneous elements such as facial landmarks and background clutter. This step is crucial for focusing the segmentation process exclusively on areas prone to wrinkle formation, thereby improving segmentation accuracy by reducing noise from non-relevant regions.

Central to our methodology is the utilization of the Unet++ architecture for deep learning-based segmentation. The Unet++ model, characterized by its encoder-decoder structure and densely nested skip pathways, facilitates the extraction and reconstruction of wrinkle features with high precision. This architecture enables our model to capture fine details and nuances present in facial wrinkles, leading to more accurate segmentation results.

Moreover, during model training, we employ a combination of dice loss and focal loss functions to effectively address the challenge of class imbalance. The dice loss function penalizes incorrect predictions and encourages accurate boundary delineation, while the focal loss function assigns higher weights to challenging wrinkle regions, thereby improving the model's ability to capture fine wrinkle details.

By integrating these components into a cohesive framework, our methodology enables accurate and robust segmentation of facial wrinkles across diverse images. In the subsequent sections of the paper, we provide detailed implementation details, experimental results, and discussions on the efficacy and implications of our methodology.

3. FINDINGS AND DISCUSSION

3.1 Facial wrinkle segmentation model

Our segmentation model is based on the Unet++ architecture, renowned for its effectiveness in semantic segmentation tasks. The Unet++ model comprises an encoder, decoder, and subnetworks interconnected through densely nested skip pathways. During the segmentation process, the encoder downsamples skin region images to extract wrinkle feature maps enriched with contextual information. These downsampled wrinkle feature maps are then propagated to corresponding layers of sub-networks through densely nested skip pathways. Subsequently, the decoder upsamples features from the corresponding sub-network layers to generate wrinkle masks, effectively delineating wrinkle regions within facial images.

In addition to the model architecture, we employ two loss functions to enhance segmentation performance: dice loss and focal loss. The dice loss function measures the similarity between the predicted segmentation mask and the ground truth mask, emphasizing overlapping regions. By penalizing false predictions, dice loss effectively mitigates the challenge of class imbalance inherent in wrinkle segmentation tasks. Mathematically, the dice loss is defined by Eq. (1),

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

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

On the other hand, focal loss is utilized to address class imbalance by assigning lower weights to non-wrinkle elements and higher weights to wrinkle elements during training. This reshaped loss function is particularly beneficial in scenarios where non-wrinkle facial elements are easily distinguishable, while wrinkles pose a greater challenge for identification. In our study, we set the weighting factor α to 0.75 and the focusing parameter γ to 2, based on empirical observations. Mathematically, focal loss is defined by Eq. (2), where pt represents the predicted probability of the correct class.

$$F(p_t) = \begin{cases} -\alpha(1-p_t)^{\gamma}\log(p_t), & y = 1 \\ -(1-\alpha)p_t^{\gamma}\log(1-p_t), & otherwise \end{cases} \quad (2)$$

Together, the Unet++ architecture and the combination of dice loss and focal loss contribute to the robustness and accuracy of our facial wrinkle segmentation model. In Fig. 2, we provide qualitative comparisons of facial wrinkle segmentation methods, demonstrating the effectiveness of our approach in accurately delineating wrinkle regions while minimizing false predictions.

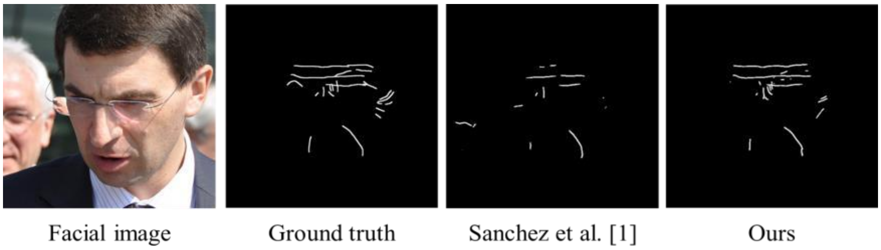

Figure 2. presents qualitative comparisons between different facial wrinkle segmentation methods.

3.2 Experiments

The qualitative assessment conducted in this experiment offers valuable insights into the visual performance of our proposed wrinkle segmentation method compared to a previous deep learning approach [1]. Fig. 2 presents a side-by-side comparison of the segmentation results obtained from both methods. Upon closer examination, it becomes evident that our method excels in accurately capturing and delineating wrinkles while effectively excluding non-wrinkle elements. Notably, the segmentation output produced by our model closely aligns with the ground truth, demonstrating a high level of fidelity in reproducing intricate wrinkle patterns. This superior performance is attributed to the robust learning capabilities of our model, which effectively leverages contextual information and intricate feature representations to distinguish wrinkles from background noise. Overall, the qualitative evaluation reaffirms the efficacy and reliability of our proposed method in achieving precise and faithful wrinkle segmentation results.

We delve into a comprehensive quantitative analysis to further elucidate the performance of various wrinkle segmentation methods, including our proposed approach. Three key performance metrics—F1 score, Intersection over Union (IoU), and pixel accuracy—are employed to provide a detailed assessment of segmentation quality.

Table 1 presents a summary of the quantitative comparison results obtained from our experiments.

Table 1. Provides a quantitative comparison of various wrinkle segmentation methods.

Segmentation Methods	F1 score ↑	IoU ↑	Pixel acc. ↑
Sanchez et al.	0.4925	0.3317	0.9919
Ours	0.5949	0.4275	0.9929

Upon analysis of the results, it is evident that our proposed method consistently outperforms the previous approach across all evaluated metrics. The higher F1 score, IoU, and pixel accuracy achieved by our method underscore its superior ability to accurately segment diverse wrinkle patterns present in facial images. This enhanced performance can be attributed to the robust architecture of our segmentation model, coupled with the effective utilization of dice loss and focal loss functions during training. By striking a balance between precision and recall, our method successfully captures subtle wrinkle details while minimizing false positives and negatives. The quantitative evaluation provides compelling evidence of the efficacy and reliability of our proposed method in achieving high-quality wrinkle segmentation results. These findings not only validate the superiority of our approach but also highlight its potential for practical applications in various domains, including dermatology, cosmetology, and age estimation.

3.3. Discussion

The discussion section provides insights into the implications of our experimental findings and further elaborates on the strengths and limitations of our proposed facial wrinkle segmentation method. Firstly, our qualitative evaluation revealed the superior performance of our method in accurately delineating wrinkles while minimizing false positives. This suggests that our model effectively learns intricate wrinkle patterns and demonstrates robustness in handling variations in wrinkle appearance across diverse facial images. The qualitative analysis also highlighted the importance of contextual information and feature representation in achieving precise wrinkle segmentation.

Moreover, our quantitative evaluation reaffirmed the effectiveness of our method, showcasing higher scores in F1 score, Intersection over Union (IoU), and pixel accuracy compared to the baseline approach. The consistent improvement across all metrics underscores the robustness and reliability of our proposed method in accurately capturing diverse wrinkle patterns present in facial images. Additionally, the successful integration of dice loss and focal loss functions contributed to mitigating the challenges posed by class imbalance, leading to enhanced segmentation performance.

However, despite the promising results, our method may still have certain limitations. For instance, the performance of the segmentation model could be influenced by factors such as lighting conditions, image quality, and variations in skin texture. Additionally, while our method demonstrates efficacy in segmenting wrinkles, further validation on larger and more diverse datasets would be beneficial to assess its generalization capability across different demographic groups and skin types. To better illustrate the advantages of our method, it is essential to compare it with a variety of advanced facial wrinkle segmentation techniques. For instance, traditional methods such as thresholding-based segmentation or edge-detection techniques often struggle with the intricacies of wrinkle patterns, particularly when dealing with fine or subtle wrinkles. These approaches generally lack the capability to handle complex texture variations and may produce suboptimal results in the presence of noise or varying lighting conditions.

The recent deep learning-based methods, such as those using different versions of the Unet architecture or other convolutional neural networks (CNNs), have shown improved performance due to their ability to learn hierarchical feature representations. However, many of these methods face challenges related to class imbalance and require extensive data augmentation or complex loss functions to achieve accurate segmentation. For example, some techniques rely heavily on data augmentation to mitigate class imbalance, which can be computationally intensive and may not fully address the problem.

4. CONCLUSION

We have presented a novel facial wrinkle segmentation method that leverages deep learning techniques to achieve accurate and robust segmentation results. Through a combination of qualitative and quantitative evaluations, we have demonstrated the superior performance of our proposed method compared to existing approaches. Our method effectively captures diverse wrinkle patterns, minimizing false positives and negatives, and showcases remarkable robustness in handling variations in wrinkle appearance.

The successful integration of dice loss and focal loss functions, along with the utilization of the Unet++ architecture, has contributed to the enhanced performance of our segmentation model. These findings underscore the potential of our method for various practical applications in dermatology, cosmetology, and age estimation.

In future work, we aim to further validate our method on larger and more diverse datasets, as well as explore avenues for improving computational efficiency and scalability. Additionally, incorporating additional contextual information and exploring novel loss functions could further enhance the performance and generalization capability of our proposed method. Overall, our study lays the groundwork for advancements in facial wrinkle segmentation and opens up avenues for future research in this domain.

REFERENCES

Dhalla, S., Maqbool, J., Mann, T. S., Gupta, A., Mittal, A., Aggarwal, P., Saluja, K., Kumar, M., & Saini, S. S. (2023). Semantic segmentation of palpebral conjunctiva using predefined deep neural architectures for anemia detection. Procedia Computer Science, 218, 328–337. https://doi.org/10.1016/j.procs.2023.01.015

Gowroju, S., Kumar, S., Aarti, & Ghimire, A. (2022). Deep Neural Network for Accurate Age Group Prediction through Pupil Using the Optimized UNet Model. Mathematical Problems in Engineering, 2022, 1–24. https://doi.org/10.1155/2022/7813701

Kim, M., & Lee, B.-D. (2021). A Simple Generic Method for Effective Boundary Extraction in Medical Image Segmentation. IEEE Access, 9, 103875–103884. https://doi.org/10.1109/ACCESS.2021.3099936

Kim, S., Yoon, H., Lee, J., & Yoo, S. (2023). Facial wrinkle segmentation using weighted deep supervision and semi-automatic labeling. Artificial Intelligence in Medicine, 145, 102679. https://doi.org/10.1016/j.artmed.2023.102679

Lboukili, I., Stamatas, G., & Descombes, X. (2023). Automatic granular and spinous epidermal cell identification and analysis on in vivo reflectance confocal microscopy images using cell morphological features. Journal of Biomedical Optics, 28(04). https://doi.org/10.1117/1.JBO.28.4.046003

Moon, C.-I., & Lee, O. (2022). Skin Microstructure Segmentation and Aging Classification Using CNN-Based Models. IEEE Access, 10, 4948–4956. https://doi.org/10.1109/ACCESS.2021.3140031

Mukasheva, A., Koishiyeva, D., Suimenbayeva, Z., Rakhmetulayeva, S., Bolshibayeva, A., & Sadikova, G. (2023). Comparison evaluation of unet-based models with noise augmentation for breast cancer segmentation on ultrasound image. Eastern-European Journal of Enterprise Technologies, 5(9 (125)), 85–97. https://doi.org/10.15587/1729-4061.2023.289044

Nader, R., Smorodin, A., De La Fourniere, N., Amouriq, Y., & Autrusseau, F. (2022). Automatic teeth segmentation on panoramic X-rays using deep neural networks. 2022 26th International Conference on Pattern Recognition (ICPR), 4299–4305. https://doi.org/10.1109/ICPR56361.2022.9956708

Yang, M.-Y., Shen, Q.-L., Xu, D.-T., Sun, X.-L., & Wu, Q.-B. (2024). Striped WriNet: Automatic wrinkle segmentation based on striped attention module. Biomedical Signal Processing and Control, 90, 105817. https://doi.org/10.1016/j.bspc.2023.105817

Yoon, H., Kim, S., Lee, J., & Yoo, S. (2023). Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis. Diagnostics, 13(11), 1894. https://doi.org/10.3390/diagnostics13111894

Zheng, Q., Purwar, A., Zhao, H., Lim, G. L., Li, L., Behera, D., Wang, Q., Tan, M., Cai, R., Werner, J., Sng, D., van Steensel, M., Lin, W., & Kot, A. C. (2022). Automatic Facial Skin Feature Detection for Everyone. https://doi.org/10.48550/ARXIV.2203.16056

Sanchez, M., Triginer, G., Ballester, C., Raad, L., & Ramon, E. (2022, December). Photorealistic Facial Wrinkles Removal. In Asian Conference on Computer Vision (pp. 117-133). Cham: Springer Nature Switzerland.