Occluded Face Recognition

Subjects: Computer Science, Artificial Intelligence

Contributor: Honglei Li , Yifan Zhang , Wenmin Wang , Shenyong Zhang , Shixiong Zhang

Occlusion in facial photos poses a significant challenge for machine detection and recognition. Consequently, occluded face recognition for camera-captured images has emerged as a prominent and widely discussed topic in computer vision. The standard face recognition methods have achieved remarkable performance in unoccluded face recognition but performed poorly when directly applied to occluded face datasets. The main reason lies in the absence of identity cues caused by occlusions. Therefore, a direct idea of recovering the occluded areas through an inpainting model has been proposed.

occluded face recognition   identity-guided inpainting   image synthesis

generative adversarial net (GAN)

1. Introduction

In recent years, occluded face recognition has become a research hotspot in computer vision. Unlike unoccluded faces, occluded faces suffer from incomplete visual components and insufficient identity cues, which lead to degradation in recognition accuracy by normal recognizors [1][2][3][4]. Inspired by the recovery mechanism of the nervous system, researchers have proposed two types of approach, i.e., occlusion-robust and occlusion-recovery.

The occlusion-robust approach attempts to improve the robustness of recognizers on occluded faces by improving the "representation". The latest work, FROM [5], proposed an end-to-end occluded face recognition model to learn the feature masks and deep occlusion-robust features simultaneously. However, compared with normal recognizers, it has weakened generalization ability over datasets with wide age and angle differences, such as the CFP-FP [6] and AgeDB-30 [7].

Unlike the occlusion-robust approach, the occlusion-recovery approach recovers the occluded regions before recognition. GAN-based inpainting methods [8][9] have remarkably improved realistic content generation. At the same time, identity-preserving inpainting models [10][11][12][13][14][15] have been demonstrated to be effective for occluded face recognition. These methods often adopt encoder-decoder-structured networks but with different identity loss during training, as **Figure 1** shows. Dolhansky et al. [10] imported identity features to preserve identity information in eye regions by L2 feature loss, as **Figure 1**b shows. Inspired by the perceptual loss [11][12][16] used identity loss which combined perceptual items and identity feature items, as **Figure 1**c shows. The perceptual item is computed with semantic features from a low-level layer of the pretrained recognizer, while the identity feature item is from the output of the top-level layer. Ge et al. [15] proposed an identity-diversity loss that combines

perceptual loss and identity-centered triplet loss to guide face recovery, which achieved state-of-the-art performance in identity preserving inpainting, as **Figure 1**d shows. Duan et al. [13] designed two-stage GAN models to deal with face completion and frontalization simultaneously. However, these methods are also limited by the challenge of preserving the inherent identity information against large occlusions. These methods often utilize incomplete datasets to learn the identity distribution with the supervision of identity and reconstruction loss functions, which makes the learned distribution deviate from its real one. Then, the decoder generates a new face from sampling the biased identity space, further enhancing the identity offset of the generated image.
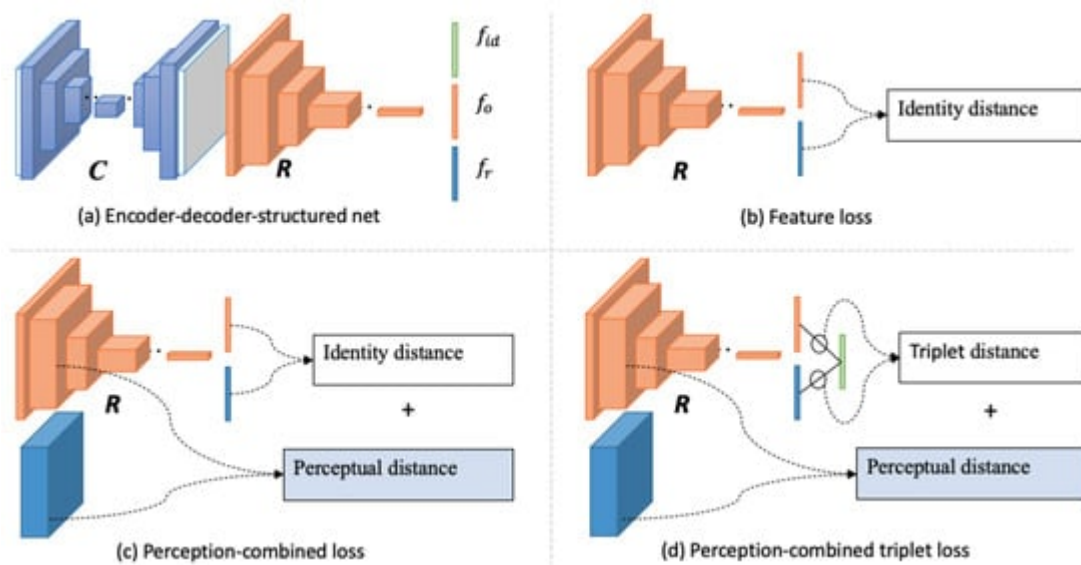

Figure 1. Encoder-decoder-structured identity-preserving inpainting networks with different identity training loss. C is an encoder-decoder-structured content inpainting network, and R is a pretrained recognizer. fid, fo, fr are identity-centered features, occlusion-recovered features, and real face features, respectively.

2. Occluded Face Recognition

Face recognition is a computer vision task that recognizes the identity among multiple face images. It is closely related to feature extraction, classification [17], and detection [18] technology. As one of the most successful practical cases, face recognition has a long history of research which has extended to various application scenarios [15][19][20]. Traditional face models are designed for unoccluded face images (see, for example, [1][2]). When they are applied directly to occluded datasets, their accuracy drops dramatically. There are two main approaches to solving the problem: occlusion-robust and occlusion-recovery.

The occlusion-robust approach reduces the accuracy drop by improving the robustness of recognizers on occluded faces. One idea is to improve the "representation". Refs. [21][22][23] report various kinds of representation methods for facial features. The latest work called FROM [5] is an end-to-end occluded face recognition model to learn the feature masks and deep occlusion-robust features simultaneously and achieved the SOTA result on the occluded LFW dataset.

Unlike the occlusion-robust approach, the occlusion-recovery approach recovers the occluded facial regions and then performs recognition on the recovered faces. Ge et al. [15] proposed an identity-diversity inpainting network to facilitate occluded face recognition. It improved the recovery step by integrating GAN with a novel CNN network, which used identity-centered features as supervision to enable the inpainted faces to cluster towards their identity centers. In [14], occlusions were removed with a CNN-based deep inpainting network. However, these methods are also limited by the challenge of preserving the inherent identity information against large occlusions. The core reason lies in the insufficient transformation of identity information. So, if we can improve the identity information transformation in the inpainting phase, we will further improve the performance of occluded face recognition.

3. Identity-Preserving Face Inpainting

A simple approach for face inpainting is to borrow general deep learning inpainting methods directly, which are good at rebuilding the overall structure of the face. For example, generative inpainting methods [9][24] involve the design of attention layers to improve the global structure consistency and fidelity and have performed well in face inpainting. Although these methods have been shown to maintain the consistency of facial structure, they showed limited improvement in occluded face recognition. So, some researchers have turned their attention to identity-preserving face inpainting.

Identity-preserving face inpainting attempts to perceive the identity information from the uncorrupted region. Some attempts, e.g., [14][15][25], imported identity loss to solve the problem and were demonstrated to be effective for occluded face recognition, but not significantly. For example, Ge et al. [15] proposed an identity-preserving face completion model that combined a CNN network and a third recognizer player to complete identity-diversity inpainting. It was designed explicitly for occluded face recognition but failed to improve performance on large-size occlusions. The main reason is that the traditional encoder-decoder network trained on occluded datasets can not build real identity space, leading to a prominent identity offset in the inpainting process. Li et al. [26] creatively combined a general inpainting network with AAD-generator [27] to solve identity-guided inpainting tasks, regenerating missing content from a pretrained identity distribution. However, there is still a certain distance in style and structure between the generated face and the ground truth face. Although an additional Poisson blending module is used to repair the style difference, the structure bias cannot be erased.

4. Normalization Layers

GAN is powerful in generating photo-realistic results based on distribution sampling. There have been broad investigations of the normalization layers [26][27][28][29] in GANs to improve the prediction performance.

References

1. Wang, H.; Wang, Y.; Zhou, Z.; Ji, X.; Gong, D.; Zhou, J.; Li, Z.; Liu, W. Cosface: Large margin cosine loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 5265–5274.

2. Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4690–4699.

3. Deng, J.; Guo, J.; Yang, J.; Lattas, A.; Zafeiriou, S. Variational prototype learning for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 11906–11915.

4. Huang, Y.; Wang, Y.; Tai, Y.; Liu, X.; Shen, P.; Li, S.; Li, J.; Huang, F. Curricularface: Adaptive curriculum learning loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 5901–5910.

5. Qiu, H.; Gong, D.; Li, Z.; Liu, W.; Tao, D. End2End occluded face recognition by masking corrupted features. IEEE Trans. Pattern Anal. Mach. Intell. 2021, 44, 6939–6952.

6. Sengupta, S.; Chen, J.C.; Castillo, C.; Patel, V.M.; Chellappa, R.; Jacobs, D.W. Frontal to profile face verification in the wild. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 7–10 March 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–9.

7. Moschoglou, S.; Papaioannou, A.; Sagonas, C.; Deng, J.; Kotsia, I.; Zafeiriou, S. Agedb: The first manually collected, in-the-wild age database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 51–59.

8. Yu, J.; Lin, Z.; Yang, J.; Shen, X.; Lu, X.; Huang, T.S. Free-form image inpainting with gated convolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 4471–4480.

9. Zheng, C.; Cham, T.J.; Cai, J. Pluralistic image completion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 5–20 June 2019; pp. 1438–1447.

10. Dolhansky, B.; Ferrer, C.C. Eye in-painting with exemplar generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7902–7911.

11. Li, C.; Ge, S.; Hua, Y.; Liu, H.; Jin, X. Occluded face recognition by identity-preserving inpainting. In Cognitive Internet of Things: Frameworks, Tools and Applications; Springer: Cham, Switzerland, 2020; pp. 427–437.

12. Duan, Q.; Zhang, L. Look more into occlusion: Realistic face frontalization and recognition with boostgan. IEEE Trans. Neural Netw. Learn. Syst. 2020, 32, 214–228.

13. Duan, Q.; Zhang, L.; Gao, X. Simultaneous face completion and frontalization via mask guided two-stage GAN. IEEE Trans. Circuits Syst. Video Technol. 2021, 32, 3761–3773.

14. Din, N.U.; Javed, K.; Bae, S.; Yi, J. A novel GAN-based network for unmasking of masked face. IEEE Access 2020, 8, 44276–44287.

15. Ge, S.; Li, C.; Zhao, S.; Zeng, D. Occluded face recognition in the wild by identity-diversity inpainting. IEEE Trans. Circuits Syst. Video Technol. 2020, 30, 3387–3397.

16. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016, Proceedings, Part II 14; Springer: Cham, Switzerland, 2016; pp. 694–711.

17. Ullah, A.; Jami, A.; Aziz, M.W.; Naeem, F.; Ahmad, S.; Anwar, M.S.; Jing, W. Deep Facial Expression Recognition of facial variations using fusion of feature extraction with classification in end to end model. In Proceedings of the 2019 4th International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST), Karachi, Pakistan, 10–11 December 2019; pp. 1–6.

18. Ahmad, T.; Ahmad, S.; Rahim, A.; Shah, N. Development of a Novel Deep Convolutional Neural Network Model for Early Detection of Brain Stroke Using CT Scan Images. In Recent Advancements in Multimedia Data Processing and Security: Issues, Challenges, and Techniques; IGI Global: Hershey, PA, USA, 2023; pp. 197–229.

19. Zhang, T.; Wiliem, A.; Yang, S.; Lovell, B. Tv-gan: Generative adversarial network based thermal to visible face recognition. In Proceedings of the 2018 International Conference on Biometrics (ICB), Gold Coast, QLD, Australia, 20–23 February 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 174–181.

20. Afzal, S.; Ghani, S.; Hittawe, M.M.; Rashid, S.F.; Knio, O.M.; Hadwiger, M.; Hoteit, I. Visualization and Visual Analytics Approaches for Image and Video Datasets: A Survey. ACM Trans. Interact. Intell. Syst. 2023, 13, 1–41.

21. Qian, J.; Yang, J.; Zhang, F.; Lin, Z. Robust low-rank regularized regression for face recognition with occlusion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 23–28 June 2014; pp. 21–26.

22. Wei, X.; Li, C.T.; Lei, Z.; Yi, D.; Li, S.Z. Dynamic image-to-class warping for occluded face recognition. IEEE Trans. Inf. Forensics Secur. 2014, 9, 2035–2050.

23. Xiong, C.; Zhao, X.; Tang, D.; Jayashree, K.; Yan, S.; Kim, T.K. Conditional convolutional neural network for modality-aware face recognition. In Proceedings of the IEEE International Conference

on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3667–3675.

24. Yu, J.; Lin, Z.; Yang, J.; Shen, X.; Lu, X.; Huang, T.S. Generative image inpainting with contextual attention. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 5505–5514.

25. Mathai, J.; Masi, I.; AbdAlmageed, W. Does generative face completion help face recognition? In Proceedings of the 2019 International Conference on Biometrics (ICB), Crete, Greece, 4–7 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–8.

26. Li, H.; Wang, W.; Yu, C.; Zhang, S. SwapInpaint: Identity-specific face inpainting with identity swapping. IEEE Trans. Circuits Syst. Video Technol. 2021, 32, 4271–4281.

27. Li, L.; Bao, J.; Yang, H.; Chen, D.; Wen, F. Faceshifter: Towards high fidelity and occlusion aware face swapping. arXiv 2019, arXiv:1912.13457.

28. Park, T.; Liu, M.Y.; Wang, T.C.; Zhu, J.Y. Semantic image synthesis with spatially-adaptive normalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2337–2346.

29. Liu, H.; Wan, Z.; Huang, W.; Song, Y.; Han, X.; Liao, J. Pd-gan: Probabilistic diverse gan for image inpainting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 9371–9381.