Proceedings of CONF-SEML 2026 Symposium: Computational Analysis and Modeling in Complex Intelligent Systems
DOI: 10.54254/2755-2721/2026.GU33390

Robust Face Recognition Technology in Complex Environments for Robotic Applications

Xi Cheng

College of Computer Science, Wuhan University, Wuhan, China
2024302111244@whu.edu.cn

Abstract. In recent years, with the rapid development of intelligent robots and artificial intelligence technologies, face recognition has become the core fundamental technology for service robots, human-machine interaction robots, and security inspection robots to achieve identity perception and natural interaction. However, the actual working environment of robots often has problems such as drastic changes in lighting, frequent occlusions, diverse postures, and unstable imaging quality, which seriously affect the robustness of face recognition. Therefore, scholars at home and abroad have conducted extensive research in traditional feature design, deep learning modeling, image enhancement and restoration, etc., to improve the visual perception ability of robots in complex environments. This paper focuses on the application requirements of robots and systematically reviews the research status of robust face recognition technology in complex environments: analyzing the influence of scene environmental factors on recognition, summarizing the research progress of traditional methods and deep learning methods, summarizing the trends of low-light enhancement, multi-degradation restoration, and joint optimization of recognition, and looking forward to future directions based on real-time and embedded deployment requirements, which can provide references for related research and engineering applications.

Keywords: service robot, face recognition, complex environment, deep learning

1. Introduction

With the rapid development of service robots and intelligent interactive systems, it has become a research focus to enable robots to have reliable environmental perception and identity recognition capabilities. Among them, face recognition is widely used in scenes such as greeting robots, security inspection robots and escort robots for identity confirmation and human-computer interaction because of its non-contact and good interactive experience. However, robots usually work in open and dynamically changing environments, and often face problems such as illumination changes, occlusion interference, pose changes and motion blur. These factors will lead to significant changes in the distribution of facial features, reduce the stability and accuracy of the recognition algorithm, and affect the reliability of the system.

Face recognition in robot vision systems usually includes four main stages: face detection, face alignment, feature extraction and identity matching [1]. In different scenarios such as service,

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.GU33390

security and industry, it will be affected by low light, occlusion, low resolution and image quality fluctuation, and low light will significantly reduce the performance of face recognition [2]. Occlusion will cause the loss of key feature information, and the recognition effect will be significantly reduced when the face resolution is lower than about 32×32 pixels [3]. At the same time, robot applications have special requirements for real-time, lightweight and long-term stable operation of the system, which also brings additional constraints to the implementation of face recognition technology.

In the early research, artificial feature methods such as LBP and wavelet analysis are mostly used to enhance illumination invariance, but their generalization ability is limited in complex environments. In recent years, deep learning methods have significantly improved the feature representation ability, and combined with attention mechanism, Generative adversarial network (GAN), Vision Transformer and multi-degradation recovery model, the recognition performance in complex environments has been effectively improved. At the same time, the lightweight model and the augmentation-recognition integrated framework oriented to the real-time operation requirements of robots have gradually become a research hotspot. Although some progress has been made, problems such as complex environment adaptability, embedded deployment and multi-factor coupling modeling still need to be further studied. Based on this, this paper focuses on the application requirements of robots, systematically reviews the robust face recognition technology in complex environments, and analyzes the future development trend.

2. Robust face recognition methods based on traditional features

2.1. Feature extraction and illumination compensation

The Local Binary Pattern (LBP) is a method that is fairly accepted and employed for describing surface texture found on faces. This builds on a local structure and encoded into binary patterns based on the contrast of the pixel value with the value of its neighbors, and thus solidifying the texture descriptive feature. One of the reasons it is favorite for face recognition is the fact that it is computationally light and it can still perform with a decent accuracy under varying illumination. To pact on this deficiency, numerous researchers have developed several improved versions such as uniform LBP, multi-scale LBP, and rotation invariant LBP [4]. These positive contributions lead to stable features and improved recognition effectiveness because the sub-methods optimized representation of features, or by adding multi-scale information. All in all, LBP methods perform described very well the local texture on faces and are able to recognize faces fairly well, but only in a limited highly controlled environment.

There are studies that were also concerned with minimizing the impact of light and shadow, as a variable, by using frequency-domain analysis, as well as, illumination compensation. A method for deep face features under extreme illumination variations, is a face illumination recovery method by Hu et.al, that improves the resilience of facial representation under extreme lighting variation. In combination with face recognition, wavelet decomposition can improve the feature discrimination. All in all, preprocessing techniques and other methods, such as image normalization or histogram equalization, in conjunction with feature extraction helps improve recognition fair, despite the varying and complex illumination present in the images.

DOI: 10.54254/2755-2721/2026.GU33390

2.2. Performance and applicability in robot scenarios

Standard speculative techniques take little calculation and serve simplified needs from the hardware. They also have simplistic structures, and as a result, they tend to have consistent recognition performance on both small and medium datasets as well as fairly uncomplicated environments. Similarly, the methods possess pronounced interpretability, which underlines the ability to analyze the impact of a particular feature. Conversely, the techniques limit recognition performance from being consistent in open and variable environments. They rely on the designers' features, and because of this, when the illumination changes, when there's occlusion, or when the poses change, the techniques perform poorly.

Considerably, the robotic applications necessitate tedium and system efficiency. The traditional speculative techniques are simpler to operate on embedded devices, and as a result, they meet the real-time processing requirements to a certain extent. However, the applications have the recognition performance still limited continuity, and when the posture changes or when the illumination alters, the techniques image quality also undergoes substantial changes. As a result, the techniques are often used alongside more advanced algorithms to enhance the recognition performance.

3. Robust face recognition method based on deep learning

3.1. CNN and attention-based feature learning

Face recognition technology has been using Convolutional Neural Networks (CNN) to achieve face recognition from shallow pixel-level representation to a deep learning mechanism that extracts semantic features. Luevano et al. noted that deep architectures like ResNet can extract descriptors that are highly illumination and pose invariant [1]. Zhou et al. included LBP in a deep network and leveraged texture priors to model local deformation tolerance [5]. Hu et al. proposed a face illumination recovery method for deep face features under severe illumination variation, further improving robustness under challenging lighting conditions [6]. Ming et al. further investigated enhancement and denoising in the Fourier frequency domain and helped to retain useful identity information in degraded imaging conditions [7]. These findings confirm the ability of CNN-based techniques to improve recognition in complicated environments.

Based on this, the attention mechanism has even greater ability to improve the model to concentrate on relevant informative regions of the face and disregard irrelevant or degraded regions. Lan et al. integrated a coordinate attention module into the residual network to improve the extraction of geometric information from facial features [8]. Fredj et al. applied an adaptive fusion mechanism to allow the model to concentrate on eye regions, which are often unoccluded, mask occlusion [9]. Attention mechanisms improve feature selection which helps recognition systems to be more robust in complicated environments, in particular, in the presence of occlusion.

3.2. Loss optimization and generative restoration

Optimizing the structure and functions of a neural network can enhance the performance of a system. For example, integrating Center Loss and Softmax can increase compactness within a cluster of features for a specific identity, and increase the distance of a feature to the cluster of a different identity. A face perception system developed with neural networks by Yang and Song, and trained a neural network with a variety of face images with different illumination, and improved the

DOI: 10.54254/2755-2721/2026.GU33390

robustness of face recognition [10]. This supports the improvement of the loss function to improve the objective discriminative power of the feature space loss function.

The challenge of analyzing critically degraded face images can also be approached with generative modeling and transformer architectures. Yang et al. developed a framework for weak-light enhancement using transformer architecture and demonstrated the effective of global modeling in low-light scenario [11]. Attention mechanisms, within the context of unconstrained face recognition with mask occlusions, enable the recognition model to better focus on informative local regions of the face [3]. Joint face image restoration and frontalization for recognition and effective preprocessing of side views and severely degraded images was also addressed by Tu et al. [12]. All of these methods enhance robotic face recognition in harsh conditions.

4. Image enhancement and recognition optimization

Although deep learning has advanced facial recognition technology, there are still challenges when attempting to recognize faces in practical robotic applications. Problems, such as low illumination, motion blur, low resolution, etc., are all factors contributing to the erosion of valuable facial attributes, ultimately lowering recognition accuracy. Therefore, the simultaneous optimization of recognition and image enhancement has become a critical avenue of research.

4.1. Enhancement, recovery, and joint learning

Illumination is a critical factor in face recognition because it reduces recognition accuracy and adversely affects the performance of subsequent tasks [2]. Yang et al. developed a low-light enhancement Transformer-GAN based on global illumination modeling techniques [11]. Similarly, Ming et al. decoupled enhancement and denoising in the Fourier domain and optimized image quality for recognition-related features, thereby improving downstream face recognition performance [7]. Moreover, for unconstrained face recognition under mask occlusion, attention-based frameworks can mitigate the disruptive effects of occlusion by focusing on unoccluded and region-relevant facial components [3].

In practice, degradation factors such as blur, noise, pose deviation, and low resolution tend to coexist, which further increases the likelihood of recognition failure. Tu et al. proposed an algorithm that integrates face image restoration and frontalization, producing output images that are more compatible with recognition systems and improving facial feature representation and matching performance [12]. Compared with the traditional serial pipeline of enhancement followed by recognition, current studies pay more attention to the joint learning of enhancement and recognition. Liu et al. systematically benchmarked low-light image enhancement algorithms and demonstrated that enhancement quality can significantly affect downstream visual tasks such as face analysis [2].

4.2. Current issues and trends

Despite these achievements in this line of research, numerous challenges still remain. Firstly, training data and real robot scenes differ obviously, and this influences the overall performance of the model in generalization [1]. Second, the computational complexity of most models is high, and thus it is challenging to address the real-time constraints of embedded platforms. Third, even in severe degradation conditions, the balance between the quality of restoration and the consistency of identity is still hard to achieve, which directly influences the quality of recognition results.

DOI: 10.54254/2755-2721/2026.GU33390

Future studies can further use multi-modal information to enhance the stability of all-weather recognition and create lightweight networks to implement embedded systems. Meta-learning and domain adaptation can be useful in enhancing cross-scene generalization whereas a more integrated framework can be used to assist multi-degradation joint processing. Moreover, self-learned and unlearned can alleviate manual label reliance, and more comprehensive end-to-end optimization between enhancement and recognition could enable image enhancement to be more useful to identity discrimination, thus enhancing overall face recognition system performance in challenging settings.

5. Conclusion

In conclusion, face recognition in complex environments is an important technology for robot identity perception and natural human-robot interaction. With the continuous development of robotics, face recognition is being used more widely in service robots, companion robots, and security inspection robots. However, compared with ideal laboratory conditions, real robot environments are much more complicated. Changes in lighting, occlusion, image blur, and differences in shooting distance or angle can all affect the final recognition result. Therefore, improving the robustness of face recognition in complex environments has become an important research topic.

From the previous analysis, it can be seen that different methods have their own advantages and disadvantages. Manual feature methods are inexpensive, fast, and traditional and therefore still useful in certain tasks where only a limited amount of hardware is available. But they have poorer capability of describing facial information and the performance is not consistent enough in open situations. The recognition accuracy has been significantly increased by deep learning methods when it comes to addressing lighting changes and occlusion issues. There are also studies that have integrated image enhancement and face recognition that contributes to mitigating the adverse impact of degraded image. Simultaneously, lightweight CNN models offer a potential solution in terms of accuracy and real-time performance. Moreover, both heterogeneous recognition and multi-modal fusion techniques have demonstrated promising potential to enhance recognition in special situations. Nevertheless, these approaches continue to suffer issues including high computing cost, inadequate generalization or the absence of full-fledged engineering solutions.

Generally, there are a number of challenges associated with face recognition of robots. Complex environmental factors can be very likely to occur simultaneously and are more challenging to recognize. Embedded robotic devices also have limited computing power which limits the use of large-scale models. Moreover, models that have been trained on a given scene do not always work well when applied to a different scene. Research in the future can more concentrate on practical application requirements. Indicatively, the design of lightweight models can assist the face recognition system to be more responsive to embedded systems. In other circumstances, cross-domain learning has the power to enhance model generalization. Meanwhile, integrating multi-modal information and optimizing the entire recognition process can also enhance system robustness. As these technologies are advanced, the face recognition will be more frequently used in the real robot applications.

References

[1] Luévano, L. S., Chang, L., Méndez-Vázquez, H., Martínez-Díaz, Y., & González-Mendoza, M. (2021). A study on the performance of unconstrained very low resolution face recognition: Analyzing current trends and new research

DOI: 10.54254/2755-2721/2026.GU33390

directions. IEEE Access, 9, 75470–75493. https: //doi.org/10.1109/ACCESS.2021.3080712
[2] Liu, J., Xu, D., & Yang, W. (2021). Benchmarking low-light image enhancement and beyond. International Journal of Computer Vision, 129(2), 531–559. https: //doi.org/10.1007/s11263-020-01418-8
[3] Zhang, M., Zhang, Y., & Zhang, Q. (2023). Attention-mechanism-based models for unconstrained face recognition with mask occlusion. Electronics, 12(18), 3916. https: //doi.org/10.3390/electronics12183916
[4] Karanwal, S. (2022). Improved local binary pattern for face recognition. In L. Troiano, A. Vaccaro, N. Kesswani, I. Díaz Rodríguez, & I. Brigui (Eds.), Progresses in artificial intelligence & robotics: Algorithms & applications (pp. 86–96). Springer. https: //doi.org/10.1007/978-3-030-98531-8_9
[5] Zhou, L., Gao, M., & He, C. (2021). Study on face recognition under unconstrained conditions based on LBP and deep learning. Journal of Computational Methods in Sciences and Engineering, 21(2), 497–508. https: //doi.org/10.3233/JCM-204595
[6] Hu, H., Li, Q., Zhao, Y., & Zhang, J. (2021). Face illumination recovery for the deep learning feature under severe illumination variation. Pattern Recognition, 111, 107724. https: //doi.org/10.1016/j.patcog.2020.107724
[7] Ming, F., Wei, Z., & Zhang, J. (2024). Unsupervised low-light image enhancement in the Fourier transform domain. Applied Sciences, 14(1), 332. https: //doi.org/10.3390/app14010332
[8] Lan, J., Jiang, X., Lin, G., Zhou, X., You, S., Liao, Z., & Fan, Y. (2023). Expression recognition based on multi-regional coordinate attention residuals. IEEE Access, 11, 63863–63873. https: //doi.org/10.1109/ACCESS.2023.3285781
[9] Fredj, H. B., Bouguezzi, S., & Souani, C. (2021). Face recognition in unconstrained environment with CNN. The Visual Computer, 37(11), 4765–4776. https: //doi.org/10.1007/s00371-020-01794-9
[10] Yang, Y., & Song, X. (2022). Research on face intelligent perception technology integrating deep learning under different illumination intensities. Journal of Computing and Electronic Information Management, 1(1), 35–41. https: //doi.org/10.54097/jceim.v1i1.3166
[11] Yang, S., Zhou, D., & Cao, J. (2022). Rethinking low-light enhancement via Transformer-GAN. IEEE Signal Processing Letters, 29, 1082–1086. https: //doi.org/10.1109/LSP.2022.3167331
[12] Tu, X., Zhao, J., & Liu, Q. (2022). Joint face image restoration and frontalization for recognition. IEEE Transactions on Circuits and Systems for Video Technology, 32(3), 1285–1298. https: //doi.org/10.1109/TCSVT.2021.3078512