Proceedings of CONF-SEML 2025 Symposium: Machine Learning Theory and Applications
DOI: 10.54254/2755-2721/154/2025.TJ23128

Analysis of Face Recognition Technology From 2D to 3D in the Field of Security

Ye Sun

Big Data Management and Application, Harbin Institute of Technology, Harbin, China
2023113298@stu.hit.edu.cn

Abstract: This review discusses the development of face recognition technology in the field of security, focusing on deep learning and Two dimensions (2D), Three dimensions (3D) models and their multi-modal fusion. Among them, the recognition accuracy of the deep learning model taking Convolutional Neural Network (CNN) model as an example for static image and dynamic video can reach 91.7% and 86.7% respectively. Research on the currently widely used 2D and 3D models shows that for 2D and 3D multi-modal mixed models, the verification accuracy of natural faces can be as high as 99.74% under a strict error acceptance rate of 0.001. This technique can be applied to high-security environments such as airports, and the recognition rate based on Regularized Kernel Canonical Correlation Analysis (RKCCA) can reach 85.19%, which is 15% higher than linear Canonical Correlation Analysis (CCA). In order to solve the problems in 3D face reconstruction, a Goldstein branch method was introduced, which achieved 97.10% accuracy. Exceeds traditional techniques and reduces the risk of deception. However, although the technology has made great breakthroughs, it is still faced with large dependence on data, high computational complexity, and difficulties in the deployment of edge devices. This review highlights the development potential of face recognition in the field of security, while advocating for peaceful solutions that can reconcile accuracy, efficiency, and ethical issues.

Keywords: Face Recognition, Deep Learning, Convolutional Neural Network, 2D and 3D Models.

1. Introduction

With the development of computer vision, face recognition is widely used as a biometric technology. Through the unique physiological or behavioral characteristics of the face, it becomes a biometric-based authentication and monitoring system. It has the advantages of high recognition accuracy and fast processing speed, which can help the computer system to authenticate. In the fields of security monitoring, mobile payment, medical diagnosis, financial access control, face recognition has played an important role. Compared with traditional identity verification technologies, fingerprint recognition, face recognition, DeoxyriboNucleic Acid (DNA) recognition and other technologies are more stable in application from a physiological point of view [1]. This paper focuses on the field of security, face recognition technology is not only in the monitoring of public security real-time deployment control, but also in the enterprise or community access control to replace the traditional swipe card or password for intelligent authority control. China's Skynet system combines video surveillance and databases to help track suspects. In terms of traffic, it can also identify traffic

DOI: 10.54254/2755-2721/154/2025.TJ23128

violations and capture red light or unlicensed driving behavior. However, the accuracy of face recognition decreases due to age, posture changes, ambient shadows, and partial occlusion [2]. But it also faces security issues such as fraud, privacy breaches and data misuse, with criminals and hackers potentially committing crimes by accessing faces through security holes.

In the field of technology, face recognition has also been sufficiently developed. Research shows that deep learning, an emerging technology introduced by DeepFace and DeepID, has been widely used as a technology since 2014 [3]. Then the more common and mature is 2D technology, which can first detect the face from our daily hardware like cameras or televisions (TVS), and cut it out from the video or picture to start detection. 2D technology is divided into 4 subcategories: holistic methods, geometrical methods, local texture descriptors-based methods, and deep learning-based methods. However, 2D technology is greatly affected by the environment, and when the environmental conditions and facial appearance change, the performance of 2D technology will drop sharply. In order to reduce the impact of the environment on the recognition system, 3D technology was invented. 3D technology can capture facial features from three-dimensional multi-camera systems, remote cameras or laser devices and crop filtering, which is relatively stable in the face of environmental light posture changes, will improve the efficiency of face recognition [4]. In addition, many other studies have shown that different Artificial Neural Network (ANN) architecture models, such as Principal Component Analysis (PCA) algorithm, Deep Convolution Neural Networks (DCNN), Radial Basis Function Neural Networks (RBFNN), etc., are also used for detection and recognition [1]. The "two-route model of face recognition" proposed by Bauer and adopted by Ellis and Young researched face recognition from both neuroanatomical and cognitive approaches, connects biology and anatomy together, and becomes a widely accepted model in face processing disorders [5]. To a certain degree, these technologies have aided in the advancement of facial recognition technology and offered workable concepts for further study.

The main purpose of this thesis is to study the application and problems of face recognition technology in the field of security. This chapter introduces the current background analysis and related concepts of face recognition, and discusses the related cutting-edge technologies. The second chapter further analyses the principle of 2D and 3D face recognition technology based on deep learning architecture. The third chapter compares the application results of 2D and 3D technology, discusses the advantages and disadvantages of different technologies, and puts forward the possible development direction in the future. Chapter four summarizes and generalizes.

2. Methodology

2.1. Dataset description

With the continuous development of face recognition technology, there are many opensource datasets currently available. Labeled Faces in the Wild (LFW), a 1:1 scale assessment of faces, contains more than 13,000 faces of celebrities (as shown in Figure 1), created in 2007 by Gary B et al [6]. They cover images of different lighting conditions, expressions, and poses. It provides a rich testing environment for face recognition algorithm and solves the problem of face recognition in uncontrolled environment. The Real-world Masked Face Recognition Dataset (RMFRD), obtained from the Web using a python crawler tool, has also been widely used in the field of identity verification during the pandemic. The dataset includes 5,000 photos of 525 people with their faces and faces covered, making it the largest dataset of real faces covered in the world [7]. The IARPA Janus Benchmark(IJB), commonly used in the security field, was jointly developed by the National Institute of Standards and Technology (NIST) and IARPA and consists of three different scales (IJB-A,IJB-B, and IJB-C) focused on face recognition in unconstrained environments.(as shown in Figure 2) It is mainly taken from surveillance video, news reports, social media real security scenes, including blur, occlusion,

DOI: 10.54254/2755-2721/154/2025.TJ23128

extreme light, low resolution and other complex situations [8]. It is mainly used for identity authentication, cross-pose, cross-age, cross-illumination identification in security scenarios.

Figure 1: The samples of LFW [6]

Figure 2: The samples of IJB [8]

2.2. Proposed approach

This chapter will focus on the process and technology development of face recognition in the field of security. Starting from the currently widely used deep learning framework, an in-depth analysis of important processes and models from 2D and 3D technologies will be conducted, and the Goldstein branching method will be adopted to reduce face reconstruction problems in 3D models [9].

2.2.1. The internal steps of face recognition

The first step of face recognition is to collect the face, obtain the face image or video frame from the camera, and preprocess the image data such as noise reduction and contrast enhancement. Deep learning methods such as 'You Only Look Once for Face (YOLO-Face)' can then be used to detect the location and number of faces from the input images. Then align and standardize key points like eyes, nose, and mouth. Here can use techniques such as 5-point / 68-point face key point detection, deep learning-based pose correction, affine transformation or perspective transformation. A feature vector named signature is extracted from the captured faces. It should be noted that vectors need to be specific enough to distinguish between different facial features [4]. Finally, it is matched with the face data in the database to play the role of identity verification or recognition. (As shown in figure 3).

DOI: 10.54254/2755-2721/154/2025.TJ23128

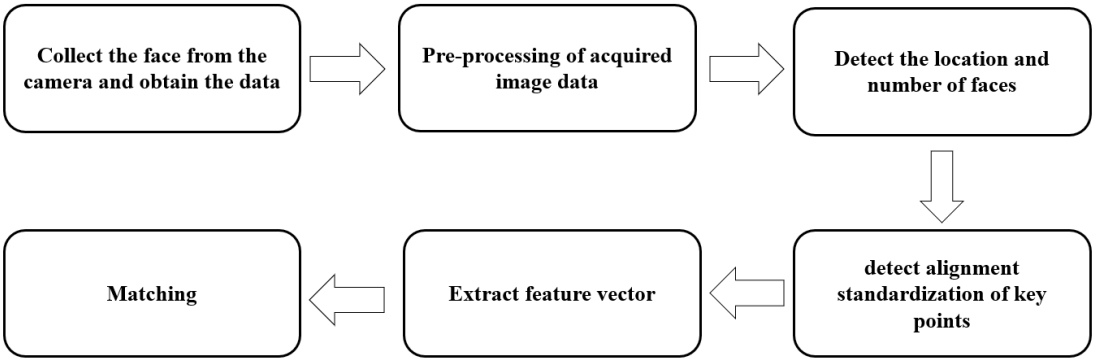

Figure 3: Steps of face recognition (picture credit: original)

2.2.2. Introduction of deep learning technology

One of the biggest dilemmas in face recognition is how to accurately recognize faces under difficult environmental conditions and changeable facial postures. Deep learning can achieve good approximation of complex functions by adding hidden layers, and achieve good results in face recognition [10]. Some studies have developed a face recognition system using Open-Source Computer Vision Library (OpenCV) in python through deep learning and achieved high accuracy [10]. There are also studies using deep Neural network architectures such as Convolutional Neural Network (CNNs) and Recurrent Neural Networks (RNNs) for evaluation and improvement. In addition, they also found that combining Variational Autoencoders (VAEs) with CNNS and RNNs can significantly increase the precision of face recognition in challenging situations [11]. These technologies have improved the security and accuracy of face recognition systems to a certain extent, and have greater application potential in the field of security. It also provides the basics for the subsequent development and application of 2D and 3D technologies.

2.2.3. D+3D recognition

Many face recognition methods over the past few decades have been based on 2D appearances, such as Eigenface and Fisherface, and are therefore sensitive to changes in lighting and pose [12]. However, some researches in the field of 2D recognition show that the maximum recognition accuracy is 99.7% by using accelerated robust feature (SURF) and scale-invariant Feature transform (SIFT) for feature extraction [13]. Some studies have also proposed some 3D face recognition algorithms and reported very high recognition rates, but 3D programs were not widely used at the beginning due to complex computation, large equipment consumption, and cumbersome preprocessing. There are studies that use a multimodal approach to combine 2D techniques and 3D models, perform PCA on intensity and distance images, and then combine the two results to get comprehensive results and accuracy, which is much higher than the accuracy formed by using 2D and 3D alone [14]. There are also studies using 2D face images as probes and 3D face data as galleries for processing, and there is good performance when employing Canonical Correlation Analysis (CCA) to learn the mapping between 2D face photos and 3D face data. Compared with the overall method, the patch-based method has significant improvement [12]. To some extent, 3D image analysis can enhance 2D face recognition ability by taking into account changes in posture, lighting, expressions, and other factors. But most 2D and 3D methods process the data separately and then merge the exchange rates and combine the same information. Therefore, in the data acquisition and processing, that need to choose the appropriate representation. For instance, it is easier to recognize facial hair in 2D photos, and the 3D segmentation procedure will be more reliable [14].

DOI: 10.54254/2755-2721/154/2025.TJ23128

2.2.4. D identification and goldstein branching method

With the popularization of intelligence, intelligent security robots are being studied and used. They can accurately obtain and recognize facial features, and feed the identified data and human behavior map back to the background database [9]. However, because of its long recognition time, it may face the problem of low recognition efficiency. Some studies have suggested the Goldstein branching method for phase unwrapping to enhance the robots' 3D face reconstruction capability and recognition efficiency in response to the challenge of intelligent security robots recognizing 3D faces. When acquiring face image data in 3D, the face data may be missing due to the influence of environment and facial posture. To reconstruct a 3D face, a four-step phase-shift stripe pattern is generated by a computer and projected onto the back and the object to be measured. When preprocessing the image data, the optimized Goldstein branch method can be used to obtain the expanded phase value and obtain the 3D image data. In the stage of feature extraction, the horizontal contour, horizontal contour and other radial curves of 3D face starting from the nose tip are obtained first, and these curves are processed by improved hierarchical matching and point distance matching methods. In the target recognition stage, thesis assign different weights to the two matching degrees and calculate the overall similarity by weighted fusion. The face recognition of 3D images is realized [9]. The 3D face reconstruction process is shown in Figure 4.

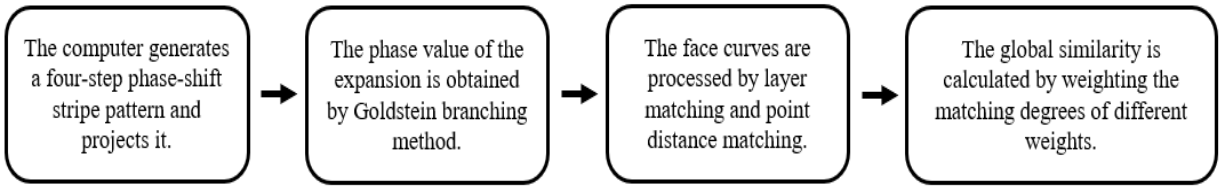

Figure 4: Steps of 3D face reconstruction (picture credit: original)

3. Results and discussion

3.1. Results analysis

Deep learning automatically extracts local and global features, such as contours, facial features, textures, through multi-layer convolution and pooling layers, avoiding the limitations of manually designed features in traditional methods. In some studies, the model using CNN architecture performed well in the recognition task, with the highest accuracy of 91.7%. In dynamic video streaming scenarios, the real-time recognition accuracy of CNN model is 86.7%, which is slightly lower than that of static images, but still meets the real-time requirements of most security scenarios, such as surveillance cameras and access control systems [10]. In addition, studies have been conducted on the multi-modal hybrid algorithm based on deep learning architecture, and the verification rates of 99.74% and 98.31% have been achieved with 0.001 error acceptance rate [15]. However, in the study of the CCA based 2D and 3D multi-modal combination model, Regularized Kernel Canonical Correlation Analysis (RKCCA) is significantly better than linear CCA, and the Rank-1 recognition rate increases from 70.37% to 85.19%. After further adopting Patch-based strategy, the recognition rate can reach 87.04%, which verifies the effectiveness of local feature matching of multi-modal data [12]. Studies using improved face reconstruction in 3D models have shown that the use of the Goldstein branch method can significantly improve the recognition accuracy, up to 97.10%. The accuracy of other methods is low, such as iterative closest point (ICP) only 95.95%, subspace pursuit (SP) only 90.42% and local binary pattern (LBP) as low as 89.73%. [9] The above results are summarized in Table 1.

DOI: 10.54254/2755-2721/154/2025.TJ23128

Table 1: Results of the models

Deep Learning (CNN)		Multimodal Fusion		Multimodal Fusion based on CCA		Reconstruct 3D Faces
Static Accuracy	91.7%	Neutral Faces	99.74%	RKCCA	85.19%	97.10%
Dynamic Accuracy	86.7%	Non-neutral Faces	98.31%	Linear CCA	70.37%	

By capturing facial depth information, 3D modeling solves the defects that 2D methods are susceptible to posture and illumination, and provides a reliable solution for in vivo detection and cross-pose recognition. And the multi-modal fusion technology can improve the intelligence and universality of face recognition in the security system.

3.2. Discussion

The biggest challenge in the face recognition system is how to improve the performance of the system in the low-resolution environment and protect the face privacy from abuse. This improves the performance of the system to some extent. However, deep learning requires a large amount of annotation data, and the face image in the security scene often has low resolution, occlusion, blurring and other problems. The 2D model can obtain face data through ordinary cameras, which has strong compatibility with existing monitoring systems and low cost. And the deep learning model is mature and supports a wide range. However, the two-dimensional model cannot distinguish the real face from the planar imitation, so it faces challenges in security performance. The 3D model developed based on the limitations of this model obtains more biometric dimensions through 3D cloud point data, which has strong anti-counterfeiting ability and high security performance. However, due to the long time required for 3D point cloud data processing and matching, 3D models require high computing power and occupy much larger storage space than 2D images.

Based on the above problems faced by different models, multi-modal fusion can be performed, combining multiple biometric features (such as 2D texture +3D geometry + infrared thermal imaging) to reduce the limitations of a single mode. Face recognition can also be combined with multi-biometric features such as gait and iris to improve reliability in face verification. Researchers can also use Neural Architecture Search (NAS) to customize security-specific models to improve reliability. Companies can develop and apply high-performance software to improve the security and reliability of facial recognition systems in the security field and Cross Match Technologies. FaceFirst is an automated facial recognition system for real-time video surveillance that is easy to use. In environments with limited resolution, it can enhance the system's recognition performance. The biggest biometrics company in the world is called Licenses. As a division of Licenses, MorphoTrak offers identity management and biometric solutions for border security, driver civil identification, law enforcement, and IT security. One of the top suppliers of biometric identity management systems, apps, and technologies for fixed, mobile, or wireless applications is Cross Match Technologies [16].

4. Conclusion

This paper provides a systematic review of the technological development of face recognition in security applications, focusing on overcoming current challenges such as lighting changes, posture diversity and data abuse, privacy breaches, etc., through deep learning architectures. The primary goal is to evaluate the efficacy of 2D, 3D, and multimodal fusion models in a deep learning architecture in improving the accuracy of high-risk security environments. CNN can be used for 2D feature extraction, 3D reconstruction by Goldstein branch method, and multi-modal fusion by

DOI: 10.54254/2755-2721/154/2025.TJ23128

RKCCA. The Goldstein branching method further refines 3D face reconstruction by minimizing surface deformation errors, which is critical for anti-spoofing applications. A large number of experimental results show that 2D-3D multi-modal fusion achieves 99.74% verification rate with 0.001 error acceptance rate, which is superior to traditional systems in airport-level security scenarios. However, challenges remain in terms of computational efficiency and edge device adaptability, especially when it comes to real-time deployment. In the future, lightweight model optimization and multimodal learning frameworks will be prioritized to address computing bottlenecks and enhance edge device compatibility. Subsequent research will focus on developing AI governance protocols to mitigate privacy breaches and ensure transparent data use. In addition, cross-domain adaptation technologies will be explored to improve performance under harsh environmental conditions, ensuring that face recognition systems are both safe and socially responsible.

References

[1] *Kasar, M.M., Bhattacharyya, D., & Kim, T.H. (2016). Face recognition using neural network: a review. International Journal of Security and Its Applications, 10(3), 81-100.*

[2] *Anwarul, S., & Dahiya, S. (2020). A comprehensive review on face recognition methods and factors affecting facial recognition accuracy. Proceedings of ICRIC 2019: Recent Innovations in Computing, 495-514.*

[3] *Wang, M., & Deng, W. (2021). Deep face recognition: A survey. Neurocomputing, 429, 215-244.*

[4] *Adjabi, I., Ouahabi, A., Benzaoui, A., & Taleb-Ahmed, A. (2020). Past, present, and future of face recognition: A review. Electronics, 9(8), 1188.*

[5] *Breen, N., Caine, D., & Coltheart, M. (2000). Models of face recognition and delusional misidentification: A critical review. Cognitive neuropsychology, 17(1-3), 55-71.*

[6] *Zheng, T., Deng, W., (2018). Cross-pose lfw: A database for studying cross-pose face recognition in unconstrained environments. Beijing University of Posts and Telecommunications, 5(7), 5.*

[7] *Wang, Z., Huang, B., Wang, G., Yi, P., & Jiang, K. (2023). Masked face recognition dataset and application. IEEE Transactions on Biometrics, Behavior, and Identity Science, 5(2), 298-304.*

[8] *Whitelam, C., Taborsky, E., Blanton, A., Maze, B., Adams, J., Miller, T., & Grother, P. (2017). Iarpa janus bench mark-b face dataset. In proceedings of the IEEE conference on computer vision and pattern recognition workshop s, 90-98.*

[9] *Wang, Z., Zhang, X., Yu, P., Duan, W., Zhu, D., & Cao, N. (2020). A new face recognition method for intelligent s ecurity. Applied Sciences, 10(3), 852.*

[10] *Teoh, K.H., Ismail, R.C., Naziri, S.Z. M., Hussin, R., Isa, M.N.M., & Basir, M.S.S.M. (2021). Face recognition and identification using deep learning approach. In Journal of Physics: Conference Series, 1755(1), 012006.*

[11] *Bein, A.S., & Williams, A. (2023). Development of deep learning algorithms for improved facial recognition in security applications. IAIC Transactions on Sustainable Digital Innovation, 5(1), 19-23.*

[12] *Yang, W., Yi, D., Lei, Z., Sang, J., & Li, S.Z. (2008). 2D–3D face matching using CCA. IEEE International Conference on Automatic Face & Gesture Recognition, 1-6.*

[13] *Gupta, S., Thakur, K., & Kumar, M. (2021). 2D-human face recognition using SIFT and SURF descriptors of face' s feature regions. The Visual Computer, 37(3), 447-456.*

[14] *Abate, A.F., Nappi, M., Riccio, D., & Sabatino, G. (2007). 2D and 3D face recognition: A survey. Pattern recognit ion letters, 28(14), 1885-1906.*

[15] *Mian, A., Bennamoun, M., & Owens, R. (2007). An efficient multimodal 2D-3D hybrid approach to automatic face recognition. IEEE transactions on pattern analysis and machine intelligence, 29(11), 1927-1943.*

[16] *Owayjan, M., Dergham, A., Haber, G., Fakih, N., Hamoush, A., & Abdo, E. (2015). Face recognition security system. In New trends in networking, computing, E-learning, systems sciences, and engineering, 343-348.*