AM1

by Teknik2 Teknik2

Submission date: 15-Mar-2024 10:22AM (UTC+0700)
Submission ID: 2320855002
File name: ancing_Biometric_Security_with_a_Broadly_Positioned_Stereo-C.pdf (3.42M)
Word count: 2521
Character count: 14782

Multi-Angle Facial Recognition: Enhancing Biometric Security with a Broadly Positioned Stereo-Camera System

Muhamad Amirul Haq[1*], *Le Nam Quoc* Huy[2], *Muhammad* Ridlwan[1], and *Ishmatun* Naila[1]

[1]Universitas Muhammadiyah Surabaya, Surabaya, Indonesia
[2]National Taiwan University of Science and Technology, Taipei, Taiwan

Abstract. This study addresses the vulnerabilities of traditional monocular camera-based face recognition systems, emphasizing the need for improved security and reliability in biometric authentication under varying environmental conditions, lighting, and human poses. To counteract the risk of spoofing attacks using masks or static images, we introduce a multi-angle stereo camera system. This system is strategically designed to capture facial imagery from multiple perspectives, thereby enhancing depth perception and spatial accuracy, crucial for high-security authentication. Employing a novel image processing approach, the study integrates a Convolutional Neural Networks (CNN) with a simple Boolean operation to differentiate the landmarks detected on each camera. This method exploits CNN's robust feature extraction capabilities and the effective usage of stereo camera, enabling precise detection and analysis of 3D facial landmarks. Such an approach significantly bolsters the system's ability to differentiate between genuine faces and deceptive representations like masks or static images. Empirical results demonstrate that the stereo camera configuration substantially improves recognition accuracy, reducing both false positives and negatives, especially in controlled spoofing scenarios. The advanced 3D facial landmark detection further reinforces the system's security. With its enhanced robustness and security, the developed system shows great potential for applications in areas requiring stringent identity verification, such as banking, public facilities, and smart home technologies.

1 Introduction

In the current digital security landscape, the reliance on conventional authentication methods, such as passwords, pin numbers, and security patterns, is ubiquitous. These methods, while familiar and straightforward, present significant challenges in terms of security and user convenience. For instance, about 56% of users who don't use password managers end up resetting their passwords every month due to forgetfulness [1]. The frequent necessity for users to reset passwords due to forgetfulness exemplifies a

* Corresponding author: amirulhaq@ft.um-surabaya.ac.id

© The Authors, published by EDP Sciences. 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/).

fundamental issue in digital security: striking a balance between ease of use and robust security measures.

However, the efficacy of these traditional security methods is increasingly being questioned. The vulnerability of password-based systems is highlighted by the alarming statistic that more than 24 billion passwords were compromised by cyber-attacks in 2022 [2]. This susceptibility, combined with the inherent risks of password theft and the difficulty in remembering complex passwords, underscores the inadequacies of conventional security approaches.

In response to these vulnerabilities, our research introduces an advanced solution: the application of a Stereo-Camera System to enhance biometric security, with a particular focus on facial recognition technology. This system offers a substantial improvement over traditional methods by utilizing the unique aspects of biometric data. Traditional monocular face recognition is vulnerable to spoof attacks using relatively simple digital images [3]–[5]. Employing two cameras to capture facial images from various angles, coupled with 3D Dense Face Alignment technology [6], the system ensures accurate and reliable individual identification and authentication, as it can differentiate between real faces and digital images. This method not only elevates security by minimizing the likelihood of spoofing attacks but also alleviates the cognitive load associated with memorizing complex passwords.

The essence of our study is to establish the effectiveness of the Stereo-Camera System in differentiating between genuine and fraudulent facial recognition attempts. Our research demonstrates that this system can significantly enhance the security of biometric authentication processes, while simultaneously maintaining user convenience. This advancement in mitigating the risks associated with password-based systems and providing a more secure, user-friendly alternative represents a significant contribution to the field of digital security. It underscores the growing relevance of biometric technologies in overcoming prevailing security challenges and sets a new benchmark for secure and efficient user authentication (**Fig. 1**).

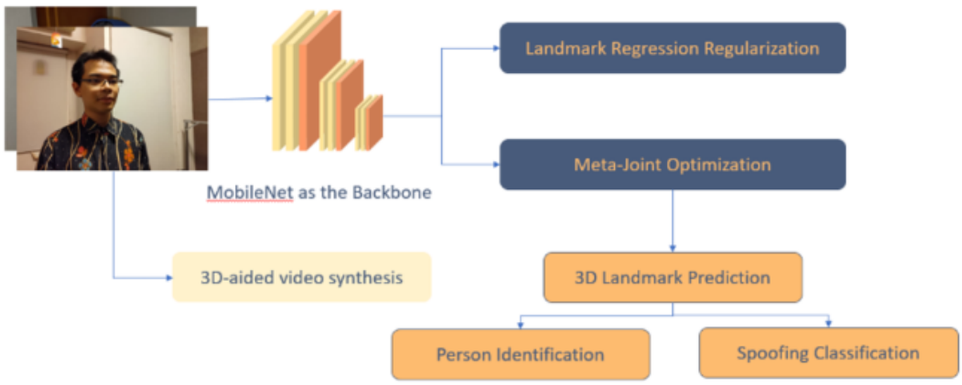


Fig. 1. Overview of the stereo-camera face recognition model architecture

2 Method

2.1 Overview

This study focuses on improving the security of face recognition using a Stereo-Camera System. The method involves using 3D Dense Face Alignment (3DDFAv2) [6] and a stereo camera setup. This setup uses two cameras to take pictures of a person's face from two different angles at the same time. Both images are processed to get 3D facial landmarks, which is also used for the person identification. If the landmarks from the images of the stereo cameras are too similar, it suggests that the face might not be real. The experiment has two parts: 'True' and 'Fake'. In the 'True' part, we take pictures of a real person from two angles at the same time. In the 'Fake' part, we use a 2D image to test if the system can detect a fake face. This approach helps us test how well the Stereo-Camera System can tell the difference between real and fake faces (**Fig. 2**).

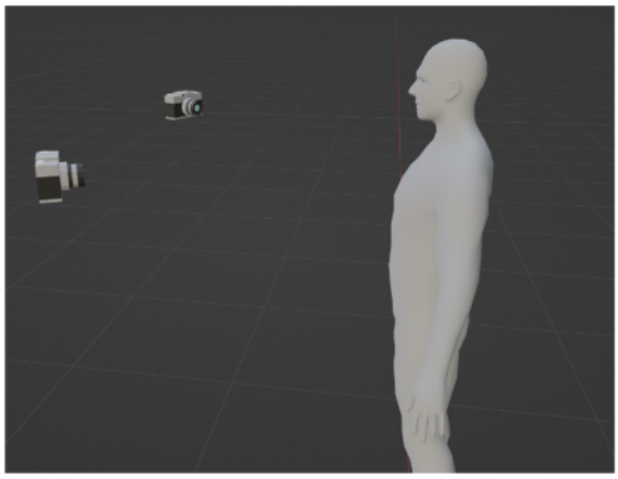


Fig. 2. Hardware acquisition system setup

2.2 Person identification and spoofing classification

The person identification system uses a 3DDFAv2 to generate the 3D face landmarks of the person and Dlib [7] face recognition framework library. 3DDFAv2 comprises of Landmark Regression Regularization, Meta-joint Optimization, and 3D Aided Video Synthesis. The Landmark Regression Regularization is a pivotal aspect of this framework, where 2D sparse landmarks are treated as an auxiliary regression task. This approach involves adding an extra landmark-regression task to the global pooling layer, trained using an L2 loss function. By flattening 68 2D landmarks into a 136-dimensional vector, the algorithm effectively emphasizes the significance of each landmark in the regression process. Meanwhile, Meta-Joint Optimization optimizes the joint distribution of parameters such as pose, expression, and illumination. This is achieved through the implementation of Vertex Distance Cost (VDC) and Weighted Parameter Distance Cost (WPDC). VDC focuses on minimizing the vertex distances between the fitted 3D model and the ground truth. In contrast, WPDC assigns varying weights to each parameter

based on their importance, thus refining the accuracy of the 3D face reconstruction. Furthermore, the 3DDFAv2 algorithm integrates a novel 3D Aided Short-Video-Synthesis approach, designed to improve the stability of 3D face alignment in video applications. This strategy addresses the common issue of random jittering in predictions, a notable challenge in existing methods. By transforming a single still image into several adjacent frames, forming a short synthetic video in a mini-batch, the algorithm models various dynamics such as noise, motion blur, in-plane rotation, and out-of-plane face movement.

All three of these components are used to generate 3D landmarks of a person's face. The resulting 3D face landmarks can then be passed on to the Person Identification Network that uses Dlib and the Spoofing Classification which uses a simple Boolean operation to compare the detected. Given landmarks detected from the first camera as a set (F_1) and from the second camera as another set (F_2). The comparison can be formulated as:

$$\text{TruePerson}(F_1, F_2) = \begin{cases} \text{true} & \text{if } F_1 \cap F_2 = \emptyset \\ \text{false} & \text{otherwise} \end{cases} \quad (1)$$

2.3 Hardware setup

The hardware for the acquisition system is set as shown in Figure 2. Two cameras are set apart so that they can capture the user's faces and obtain different landmarks. The setup is flexible and does not require calibration since the spoofing classification algorithm only considers the similarities between types of landmarks within the two sets. Therefore, the proposed method can be implemented efficiently.

2.4 Experimental setup

A Windows 11 desktop computer equipped with an Intel i7 13700KF CPU, 64 GB of RAM, and an NVidia RTX 3060-12GB VRAM GPU is used to train the suggested approach. The 3DDFAv2 model is trained on AFLW2000-3D [8], Florence [9], and Menpo-3D [10] datasets. Meanwhile, for testing purposes, we retrieve images from four subjects of various genders and ages by ourselves.

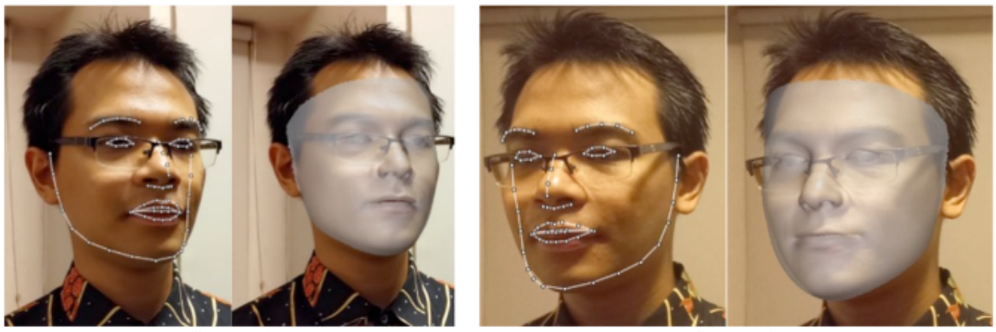
In the experimental validation of the proposed stereo camera-based facial recognition system, two meticulously designed test scenarios were employed to assess the system's efficacy in detecting spoofing attempts and its accuracy in authenticating legitimate users.

The first scenario, aimed at simulating a spoofing attack, involved the use of a tablet to display a static image of a real person's face. This setup was chosen to emulate a potential spoofing technique where an attacker might utilize a photograph or a digital representation of a legitimate user's face. The stereo cameras were then tasked with capturing the image displayed on the tablet screen, thereby testing the system's ability to discern between real human presence and digital reproductions of a face.

In contrast, the second scenario was devised to replicate a normal user authentication process. This involved the direct capture of a real person's facial images using the stereo cameras, simulating a genuine user attempting to gain access. The purpose of this scenario

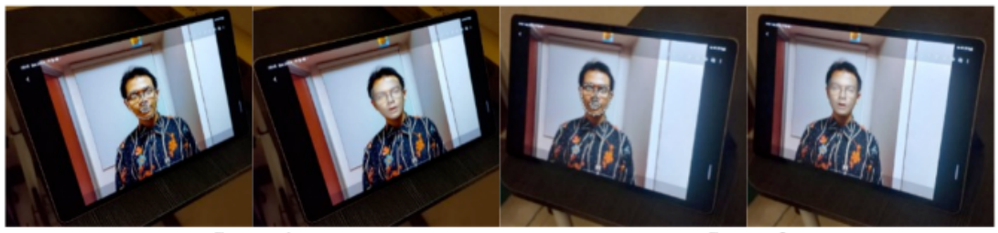
was to evaluate the system's performance under standard operating conditions, ensuring its reliability and accuracy in verifying actual users.

These two contrasting scenarios were instrumental in rigorously evaluating the system's capabilities. The spoofing simulation provided insights into the system's resilience against fraudulent access attempts, while the authentic user login simulation gauged its effectiveness in correctly identifying and granting access to legitimate users. These scenarios offered a comprehensive evaluation of the system's overall performance, encapsulating both its defensive robustness against security threats and its operational efficacy in user verification (**Fig. 3** and **Fig. 4**).



Camera 1 Camera 2

Fig. 3. "True Person" example has different landmarks from Camera 1 and 2



Camera 1 Camera 2

Fig. 4. "Fake Person" example has the same landmarks from both Camera 1 and 2, indicating a spoof attack

3 Result and discussion

The results of this study is summarized in Table 1. The method use stereo cameras and a Boolean operation for landmark comparison, demonstrates a noteworthy performance in both identifying individuals and detecting spoofing attempts. Notably, the person identification accuracy stands at 75%, while the system achieves an impressive 100% accuracy in spoofing classification.

The 75% accuracy rate in person identification, while substantial, suggests room for improvement. This figure indicates that while the system is generally reliable, there are instances where it may not correctly identify a legitimate user. These occurrences could be attributed to various factors such as variations in lighting, facial expressions, or occlusions, which are common challenges in facial recognition technologies. Future

enhancements could involve refining the person identification network that uses a more advanced model as Dlib is a lightweight model.

In contrast, the system's perfect score in spoofing classification underscores its effectiveness in distinguishing real human faces from replicas or digital images. This success can be chiefly attributed to the stereo camera setup's ability to capture distinct features from two angles, which are challenging to replicate in spoofed attempts. The Boolean comparison method proves to be particularly adept in this context, offering a simple yet fail-proof mechanism to detect discrepancies in facial landmarks indicative of spoofing.

Overall, the research presents a compelling case for exploring minimalistic approaches in biometric security systems. The findings suggest that such methods can provide substantial benefits, particularly in environments where complexity and resource constraints are key considerations. Future work could focus on addressing the limitations in person identification while maintaining the high standards achieved in spoofing detection, potentially leading to a more balanced and universally applicable facial recognition system (**Table 1**).

Table 1. Summary of the model's performance

Item	**Specification**
Desktop hardware	Intel i7-13700KF, Nvidia RTX 3060 12GB
Stereo cameras	Logitech Brio 500
Input dimension	1920 × 1080
Person identification	75%
Spoofing Classification accuracy	100%

4 Conclusion

This research introduces a novel approach to facial recognition security systems, emphasizing simplicity and effectiveness. By leveraging a stereo camera setup to capture facial images from different angles, the study's main contribution lies in utilizing a straightforward Boolean operation to compare facial landmarks. This method effectively distinguishes between real human faces and spoofed representations. The system's ability to accurately identify spoofing using this method has been demonstrated in various test scenarios, including ones that simulate digital image-based spoofing and real user authentication. The success of this research underlines the potential of combining basic logical operations with stereo imaging technology in enhancing security measures. While sophisticated systems are often the focus in the realm of biometric security, this study proves that simplicity can also lead to robust and efficient solutions.

References

1. "Global Survey: Nearly Two Thirds of People Still Rely on Memory to Recall Passwords." Accessed: Dec. 10, 2023. [Online]. Available: https://www.businesswire.com/news/home/20210428005291/en/Global-Survey-Nearly-Two-Thirds-of-People-Still-Rely-on-Memory-to-Recall-Passwords

2. "139 password statistics to help you stay safe in 2024 - Norton." Accessed: Dec. 10, 2023. [Online]. Available: https://us.norton.com/blog/privacy/password-statistics
3. F. Alqahtani, J. Banks, V. Chandran, and J. Zhang, "3D Face Tracking Using Stereo Cameras: A Review," *IEEE Access*, vol. 8, pp. 94373–94393, 2020, doi: 10.1109/ACCESS.2020.2994283.
4. Z. Li, J. Yuan, B. Jia, Y. He, and L. Xie, "An Effective Face Anti-Spoofing Method via Stereo Matching," *IEEE Signal Processing Letters*, vol. 28, pp. 847–851, 2021, doi: 10.1109/LSP.2021.3072284.
5. Z. Bai, Z. Cui, J. A. Rahim, X. Liu, and P. Tan, "Deep Facial Non-Rigid Multi-View Stereo," presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 5850–5860. Accessed: Jan. 02, 2024. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2020/html/Bai_Deep_Facial_Non-Rigid_Multi-View_Stereo_CVPR_2020_paper.html
6. J. Guo, X. Zhu, Y. Yang, F. Yang, Z. Lei, and S. Z. Li, "Towards Fast, Accurate and Stable 3D Dense Face Alignment." arXiv, Feb. 07, 2021. Accessed: Dec. 10, 2023. [Online]. Available: http://arxiv.org/abs/2009.09960
7. D. E. King, "Dlib-ml: A Machine Learning Toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.
8. X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li, "Face Alignment Across Large Poses: A 3D Solution," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 2016, pp. 146–155. doi: 10.1109/CVPR.2016.23.
9. A. D. Bagdanov, A. Del Bimbo, and I. Masi, "The florence 2D/3D hybrid face dataset," in *Proceedings of the 2011 joint ACM workshop on Human gesture and behavior understanding*, in J-HGBU '11. New York, NY, USA: Association for Computing Machinery, Dec. 2011, pp. 79–80. doi: 10.1145/2072572.2072597.
10. J. Deng *et al.*, "The Menpo Benchmark for Multi-pose 2D and 3D Facial Landmark Localisation and Tracking," *Int J Comput Vis*, vol. 127, no. 6, pp. 599–624, Jun. 2019, doi: 10.1007/s11263-018-1134-y.

AM1

ORIGINALITY REPORT

22%	15%	12%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	ebook.unimma.ac.id Internet Source	10%
2	Muhamad Amirul Haq, Le Nam Quoc Huy, Muhammad Ridlwan, Ishmatun Naila. "Leveraging Self-Attention Mechanism for Deep Learning in Hand-Gesture Recognition System", E3S Web of Conferences, 2024 Publication	5%
3	arxiv.org Internet Source	2%
4	Budi Arif Dermawan, Nani Awalia, Aries Suharso, Anis Fitri Nur Masruriyah. "The Identification of Early Blight Disease on Tomato Leaves Utilizing DenseNet Based on Transfer Learning", E3S Web of Conferences, 2024 Publication	2%
5	Rahmat Hidayat, Andi Adriansyah, Febi Kurniawan. "Development of Ceramic Decorative Rotary Tool Technology Based on the Internet of Things as a Learning Media to Support Creative Industries", E3S Web of Conferences, 2024 Publication	1%
6	"Computer Vision – ECCV 2020", Springer Science and Business Media LLC, 2020 Publication	1%
7	web.archive.org Internet Source	<1%

8	Armi Susandi, Aristiyo R. Wijaya, Mustafid Ihsan. "Challenges and Opportunities of Implementing Marine Ecological Carrying Capacity Index in the Blue Economy: A Case Study of Coastal Communities in East Nusa Tenggara Province", BIO Web of Conferences, 2024 Publication	<1%
9	Submitted to Christian University of Maranatha Student Paper	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		