



Jiawei Xu¹, Lu Leng^{1,*} and Byung-Gyu Kim^{2,*}

- ¹ Key Laboratory of Jiangxi Province for Image Processing and Pattern Recognition,
- Nanchang Hangkong University, Nanchang 330063, China; 2216083500017@stu.nchu.edu.cn
 ² Department of AI Engineering, Sookmyung Women's University, Seoul 04310, Republic of Korea
- Correspondence: leng@nchu.edu.cn (L.L.); bg.kim@sookmyung.ac.kr (B.-G.K.); Tel.: +86-791-8645-3251 (L.L.)

Abstract: At present, COVID-19 is posing a serious threat to global human health. The features of hand veins in infrared environments have many advantages, including non-contact acquisition, security, privacy, etc., which can remarkably reduce the risks of COVID-19. Therefore, this paper builds an interactive system, which can recognize hand gestures and track hands for palmvein recognition in infrared environments. The gesture contours are extracted and input into an improved convolutional neural network for gesture recognition. The hand is tracked based on key point detection. Because the hand gesture commands are randomly generated and the hand vein features are extracted from the infrared environment, the anti-counterfeiting performance is obviously improved. In addition, hand tracking is conducted after gesture recognition, which prevents the escape of the hand from the camera view range, so it ensures that the hand used for palmvein recognition is identical to the hand used during gesture recognition. The experimental results show that the proposed gesture recognition method performs satisfactorily on our dataset, and the hand tracking method has good robustness.

Keywords: infrared environment; hand gesture recognition; hand tracking; palmvein recognition

1. Introduction

The applications of biometrics are becoming more and more extensive [1]. However, due to the impact of the COVID-19, some biometric modalities are facing many challenges [2]. For example, traditional applications of face recognition require users to remove masks [3]. Some biometric applications are less affected by COVID-19, such as non-contact hand-based biometrics [4,5], which is a popular modality due to its many obvious advantages [6,7]:

- Abundant discriminative information: The hand area is large and contains rich discriminative information with a high recognition accuracy;
- Few restrictions: The hand image can still be clearly recognized in low resolution, and even if the palm or fingers are slightly abraded, the accuracy and performance are still satisfactory;
- Low cost: A low-resolution camera can meet the resolution requirement;
- Strong privacy: It is not easy to collect a user's complete hand information without his/her authorization;
- High user acceptance: Hand features can be collected in non-contact mode, thus avoiding many problems associated with contact collection.

Palmvein recognition is an important hand biometric modality used in infrared environments [8,9], which has more advantages [10]:

- Small illumination effect: In the infrared environment, the interference of the illumination difference is smaller than that in the visible environment;
- Strong liveness detection: The infrared camera can take images of human veins, which can pass through the skin tissues of the human body for liveness detection;



Citation: Xu, J.; Leng, L.; Kim, B.-G. Gesture Recognition and Hand Tracking for Anti-Counterfeit Palmvein Recognition. *Appl. Sci.* 2023, 13, 11795. https://doi.org/ 10.3390/app132111795

Academic Editor: Andrea Prati

Received: 30 September 2023 Revised: 22 October 2023 Accepted: 26 October 2023 Published: 28 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

- Difficult abrasion and forgery: As the veins of the human body belong to the internal physiological characteristics, they are usually hidden under the epidermis of the human body, and there is basically no abrasion. Therefore, it is very difficult to forge veins;
- Few restrictions: The recognition accuracy is less affected by the scars and dirts on the hand surface.

At present, most studies are focused on improving the accuracy and robustness of palmvein recognition, but do not attach enough importance to anti-counterfeiting [11,12]. This paper aims to further improve anti-counterfeiting during palmvein recognition. We build an interactive system, which can recognize hand gestures and track hands for palmvein recognition in infrared environments. Usually, illegal users try to steal genuine users' biometric images to impersonate their identities [13,14]. If the hand is required to make an "unpredictable" gesture, an illegal user cannot use a simple fake palmvein tool to pass the authentication test, because a simple tool cannot make a gesture like a real hand. Even if an illegal user can make a tool that has a certain hand gesture, he/she cannot predict the gestures commanded during authentication. Thus, this paper further improves the palmvein anti-counterfeiting method.

The contributions of this paper are summarized as follows:

- A hand gesture recognition algorithm is developed in infrared environment. Firstly, the hand gesture contours are extracted from infrared gesture images. Then, a deep learning model is used to recognize the hand gestures from these contours;
- A hand tracking algorithm is developed in an infrared environment, which is based on the detection of key points. Hand tracking is conducted after gesture recognition, which prevents the escape of the hand from the camera view range, so it ensures that the hand used for palmvein recognition is identical to the hand used during gesture recognition.

The rest of this paper is organized as follows: The works related to gesture recognition and hand tracking are reviewed in Section 2. The proposed methods for gesture recognition and hand tracking ares specified in Section 3. The experimental results are provided in Section 4. In Section 5, we conclude this work and elaborate on future improvement plans.

2. Related Works

Hand gesture recognition in an infrared environment has many advantages compared with methods using a visible environment. The hand gesture recognition in this paper is conducted in an infrared environment. Some researchers have used both a visible light environment and an infrared environment for hand gesture recognition. In 2014, Erden and Çetin proposed a multi-mode hand gesture detection and recognition system [15]. They used both infrared and visible environment information. In 2021, Yu et al. used reflected visible and infrared light signals for hand gesture recognition [16]. The simultaneous use of visible and infrared environments can enhance the system's robustness. However, sometimes, the two-camera mode results in difficulties with system development and additional costs.

Recently, as devices for collecting hand gesture information, Microsoft Kinect (MK) and the Leap Motion Controller (LMC) are becoming popular [17,18]. Kumar et al. proposed a multi-sensor fusion method based on the Hidden Markov Model (HMM) and Bidirectional Long Short-Term Memory Neural Network (BLSTM-NN), which combines MK and LMC to improve the performance of hand gesture recognition [19]. Special devices are not always available. Typically, a camera (a visible or infrared camera) is more cost-effective and has higher user acceptance.

In recent years, deep learning has been widely used [20,21]. As a result, most advanced gesture recognition methods are based on deep learning. In 2021, Mujahid et al. proposed a lightweight hand gesture recognition model based on YOLO (You Only Look Once) v3 and DarkNet-53 [22]. Qi et al. proposed the Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) for multiple hand gesture classification [23]. With the advent of

the big data era, deep learning methods commonly have higher accuracy than traditional methods [24].

Typically, fusion can effectively improve the accuracy [25] and achieve higher anticounterfeiting, high versatility, and strong robustness [26,27]. In 2022, Sahoo et al. proposed an end-to-end fine-tuning method of the pre-trained Convolutional Neural Network (CNN) model with score-level fusion for hand gesture recognition [28].

The above works are summarized in Table 1.

Table 1. Methods of hand gesture recognition.

Ref. and Year	Data	Method
[15] 2014	Infrared + Visible light	Winner-take-all Hash
[16] 2021	Infrared + Visible light	K-nearest-neighbors
[17] 2017	MK	Dynamic Time Warping
[18] 2019	LMC	Support Vector Machine
[19] 2017	MK + LMC	HMM + BLSTM-NN
[22] 2021	Visible light	YOLO v3 + DarkNet-53
[23] 2021	LMC	LSTM-RNN
[28] 2022	МК	CNN with score-level fusion

The system designed in this paper focuses on improving the anti-counterfeiting performance for subsequent palmvein recognition. It must track the hand to ensure that the hand used for palmvein recognition is identical to the hand used during gesture recognition. Hand tracking is mostly considered together with hand gesture recognition [29]. In 2013, Kulshreshth et al. described a framework of real-time finger tracking technology using MK as an input device [30]. After that, some hand tracking studies were based on the LMC, such as Houston et al.'s method [31] and Ovur et al.'s method [32].

In addition to MK and LMC, some other special sensors have also been used for hand tracking. In 2019, Nonnarit et al. proposed a virtual reality interactive interface for hand tracking based on Magnetic, Angular-Rate, & Gravity (MARG) sensors [33]. They used a set of infrared video cameras and MARG modules embedded in the users' gloves. Santoni et al. proposed a hand tracking system based on magnetic positioning [34]. They installed a magnetic node on each fingertip and two magnetic nodes on the back of their hands. They used a fixed array of receiving coils to detect magnetic fields and inferred the position and direction of each magnetic node from them.

In addition to using special equipment, some researchers used videos captured with cameras for hand tracking. Mueller et al. combined CNNs with kinematic 3D hand models to solve the real-time 3D hand tracking problem based on monocular RGB sequences [35]. Their method is robust to occlusion and different camera viewpoints. Han et al. proposed a real-time hand tracking system that can generate accurate and low jitter 3D hand movements using four fisheye monochrome cameras [36]. The system leveraged a neural network architecture to detect hands and estimate the positions of key points in the hands.

This paper focuses on hand gesture recognition in an infrared environment, rather than in a visible light environment. In addition, hand tracking is used as a bridge to connect hand gesture recognition and palmvein recognition. The proposed method can improve the security and anti-counterfeiting of palmvein recognition remarkably.

3. Methodology

Our work can be summarized as shown in Figure 1. It is worth noting that the purpose of this paper is to improve the anti-counterfeiting characteristics of palmvein recognition through gesture recognition and hand tracking, rather than to propose a new palmvein recognition method.



Figure 1. Our work.

This paper designs and implements an interactive system with two functions, namely hand gesture recognition and hand tracking. The system obtains real-time video from an infrared camera and provides a random gesture instruction to an user. The user needs to make a corresponding hand gesture in front of the infrared camera according to the instructions. After the system recognizes the correct hand gesture of the user, it will call the hand tracking function. During the operation of the hand tracking function, the user's hand is required to be within the camera view range until palmvein recognition is completed. Otherwise, the user is regarded as an illegal user by the system, who attempts to impersonate someone else's identity using a forged palmvein.

For the gesture recognition function, this paper obtains a network model after data acquisition, pre-processing, gesture contour extraction, and training. Through this network, the system can recognize whether the user's gestures comply with the instructions. For the hand tracking function, the hand key points are detected in each frame. The specific contents mentioned above are detailed in the following sections.

3.1. Data Acquisition

The infrared camera used for image acquisition is shown in Figure 2. It consists of a USB plug, a data cable, and a camera equipped with two rows of infrared fill-in lights. The USB plug and data cable are used to connect the camera to the computer. Two rows of infrared fill-in lights are used to provide an infrared environment.



Figure 2. Infrared camera used for image acquisition.

The dataset includes 10 gestures, and 100 images are collected for each gesture. The 10 hand gesture types are shown in Figure 3. Detailed descriptions of the gestures are shown in Table 2.



Figure 3. Ten hand gesture types. (**a**) Only straighten the index finger. (**b**) Only straighten the index finger and the middle finger. (**c**) Only bend the thumb and the little finger. (**d**) Only bend the thumb. (**e**) Five fingers straight and separated. (**f**) Five fingers straight and together. (**g**) Only bend the thumb and the middle finger. (**h**) Only bend the thumb and the ring finger. (**i**) Only bend the thumb and the index finger. (**j**) Only straighten the index finger and the little finger.

Contract	Fingers					Distance of Firecore
Gesture	Thumb	Index	Middle	Ring	Little	Distance of Fingers
а	В	S	В	В	В	
b	В	S	S	В	В	
С	В	S	S	S	В	
d	В	S	S	S	S	
e	S	S	S	S	S	Separated
f	S	S	S	S	S	Together
g	В	S	В	S	S	_
ĥ	В	S	S	В	S	
i	В	В	S	S	S	
j	В	S	В	В	S	

Table 2. The descriptions of the gestures. "S" means straight. "B" means bent.

3.2. Pre-Processing

When a hand gesture image is captured in the infrared environment, a user's hand is usually placed in the middle of the shooting area. In order to reduce the background on the left and right sides, the original 640×480 images are cropped to 480×480 .

Each image is rotated nine times at a random angle to augment the dataset. In addition, the angle of the user's hand placement can be flexible and unfixed. Each rotation angle is between -30° and 30° . The rotated image is added into the dataset. In addition, all images are flipped horizontally and added into the dataset. All captured gesture images are right-hand gestures, and their flipped images can be seen as left-hand gestures. In this way, the gesture recognition model can recognize both left and right hand gestures simultaneously. The augmented dataset includes 20,000 hand gesture images in the infrared environment (10 classes, 2000 images for each class).

3.3. Gesture Contour Extraction

The gesture contour extraction method extracts hand gesture contours from infrared images with pure backgrounds. Firstly, the original images are binarized. There is a bright spot on the background caused by the infrared fill-in light, so the overall brightness of the images is uneven. Therefore, adaptive threshold binarization is used. The adaptive threshold binarization is defined as:

$$f_s(n) = \sum_{i=0}^{s-1} p_{n-i}$$
(1)

$$T(n) = \begin{cases} 0, & \text{if } p_n > \frac{f_s(n)}{s} - c\\ maxValue, & \text{otherwise} \end{cases}$$
(2)

n is the center pixel of the sliding window, and its gray value is p_n . $f_s(n)$ represents the sum of the gray values of *s* pixels near *n*. T(n) is the binary result of *n*. c = 2, $s = 11 \times 11$, maxValue = 255.

Then, the median filter is used to smooth the binary images. The median filter is defined as:

$$g(x, y) = Med\{f(x - i, y - j)\}, (i, j) \in S$$
(3)

g(x, y) and f(x, y) are gray values. *Med*{f(x - i, y - j)} represents the median from the set of f(x - i, y - j). The size of the filter window *S* is 9 × 9. The smoothed images are gesture contours.

3.4. Network Architecture

Supported by a large amount of trainable data, deep learning methods commonly perform better than traditional methods [37,38]. ResNet-34 is a commonly used deep learning model for image classification tasks [39]. It solves the vanishing gradient and exploding gradient problems by introducing residual blocks [40]. Considering the good performance and low complexity of ResNet-34, this paper uses ResNet-34 as the backbone network.

The channel attention mechanism enables the model to learn the channel weights, which emphasizes the significant features with large weights while suppressing the unimportant features with small weights; therefore, it improves the model's performance [41]. The Squeeze-and-Excitation Network (SENet) is a commonly used channel attention mechanism [42]. The core part of the network, namely the Squeeze-and-Excitation (SE) block, is shown in Figure 4. The block consists of a global average pooling layer, two full connection (FC) layers, and two activation functions (ReLU and Sigmoid).



Figure 4. SE block.

This paper embeds the SE block in ResNet-34, as SE-ResNet-34, to enable the ResNet-34 to learn channel weights to improve the model's performance. The architecture of SE-ResNet-34 is shown in Figure 5. In this paper, the SE block is independent of the residual block, rather than being embedded in the residual block. We embed the SE block after each residual block, so that the network can learn the channel weights of the output features of each residual block. Due to there being only 10 types of gestures in the dataset, the output dimension of the last layer of the network is changed to 10.



Figure 5. Architecture of SE-ResNet-34.

3.5. Hand Tracking

The purpose of hand tracking is to prevent the hand from escaping, so as to ensure that the user's hand is within the camera view range at all times. This can avoid the situation where illegal users use their own hands for gesture recognition and fake palmveins for palmvein recognition, which ensures that the hand used for palmvein recognition is identical to the hand used for gesture recognition.

Google's MediaPipe model can detect 21 hand key points and locate them [43], which is used for hand tracking. When the key points can be successfully detected, the hand is considered to be successfully detected. Then, hand detection is conducted in each frame for hand tracking. As long as the hand is not detected in two consecutive frames, it is considered that the hand has escaped from the camera view range. The robustness of this algorithm is excellent [44,45]. Even if some fingers exceed the infrared camera view range, the algorithm can still work. In other words, the system can tolerate a small part of the hand leaving the camera view range.

3.6. Interactive System

The PyQt5 framework is used to build the interactive system [46,47]. The interactive system includes two image display windows, two buttons, and some text. The first image display window shows the first image of a hand gesture class randomly selected by the system from the dataset. The second image display window shows the real-time video captured by the infrared camera. In order to avoid the second image display window being stuck when the system is recognizing the gesture, two threads are added into the interactive system, which are responsible for hand gesture recognition and hand tracking, respectively. They can return Signal 0 or Signal 1 to the interactive system.

When the hand gesture recognition starts, Thread 1 is called. The current frame in the video stream is intercepted, and put into the trained network for prediction. The gesture class of the prediction result is compared with the gesture class selected by the system. When they are the same, Signal 1 is returned, which confirms that the user's hand gesture is right. If they are different, the above steps are repeated. After the above steps have been repeated five times, when Signal 1 is still not returned, the user's hand gesture is wrong and Signal 0 is returned.

After the hand gesture recognition is correct, the system prompts the user to open the palm and fingers to prepare for palmvein acquisition. At the same time, Thread 2 is called for hand tracking. When the hand can be continuously detected, Signal 1 is always returned. When the hand is not detected, Signal 0 is returned and the user is prompted by the interactive system that his/her hand has left from the infrared camera view range.

4. Results

4.1. Gesture Contour Extraction

Figure 6 shows the hand contour extraction results. The first image in the dataset is taken as an example. After the original image has undergone adaptive threshold binarization, the contours of the hand can be clearly seen on the image. However, there is still some noise left in the image at this time. After being smoothed through median filtering, this noise is removed. At this point, only the contours of the hand are retained in the image. Inputting these images, which only retain the contours of the hand, into the network can enable the network to better learn gesture features.

4.2. Accuracy

The method used in this paper is compared with some typical deep learning networks, including the typical CNN, Visual Geometry Group (VGG) [48], GoogLeNet [49], ResNet-34, typical CNN with embedded SE block (SE-CNN), three versions of ConvNeXt [50], FasterNet [51], and ResNet-18 with embedded SE block (SE-ResNet-18).

In order to better compare the learning ability of the networks for our dataset, some parameters are set in advance. The setting of the initial learning rate requires the consideration of the model's complexity and dataset size. After hyperparameter tuning, the initial learning rate is set to 0.0001. The optimization algorithm is Adam. The loss function is cross entropy. The dataset is divided into a training set and a test set with a ratio of 8:2. The number of epochs is 20.

Table 3 shows the accuracy results of the networks. The accuracy of our network is the highest. Therefore, our network is selected as the hand gesture recognition network.



Figure 6. Gesture contour extraction. (a) Adaptive threshold binarization. (b) Median filtering.

Table 3. Accuracy comparison of networks.

Network	Accuracy		
CNN	82.175%		
VGG-16 [48]	92.675%		
GoogLeNet [49]	89.800%		
ResNet-34 [40]	91.725%		
SE-CNN	90.950%		
ConvNeXt-T [50]	96.100%		
ConvNeXt-S [50]	98.625%		
ConvNeXt-B [50]	99.300%		
FasterNet [51]	97.775%		
SE-ResNet-18	98.175%		
SE-ResNet-34 (Ours)	99.425%		

Bold represents the best accuracy and its corresponding network.

4.3. System Implementation

The initial interface of the implemented interactive system is shown in Figure 7. In this paper, an infrared fill-in light is used to provide an infrared environment, so, in the real-time captured image on the right in the initial interface, a bright spot can be seen in the center of the image. When a user's palmvein is captured, it is necessary to control the distance between the hand and the camera to ensure that the entire palm can be captured.



Figure 7. The initial interface.

After the "start" button has been clicked, a hand gesture sample is randomly selected by the system and displayed, as shown in Figure 8. All hand gesture samples are the first images from all gesture classes in the dataset.



Figure 8. Identifying.

Then, the user starts to display his/her hand gesture in front of the camera, and the system starts to recognize the gesture. If the user does not make the correct hand gesture after a period of time, the interface will prompt "Wrong gesture". And, the interface will prompt the user to restart by clicking the "Start" button, as shown in Figure 9.





If the system recognizes that the user's hand gesture is correct, it will prompt the user with "Open your hand" (prepare for palmvein acquisition), as shown in Figure 10.



Figure 10. The hand gesture is correct.

Then, the system starts to track the hand. If no hand is detected in two consecutive frames, the words "Do not leave" will appear on the interface, as shown in Figure 11. If the user wants to continue to use the system, he/she needs to click the "Start" button again.



Figure 11. Hand escape.

If the hand does not escape from the camera view range and the user places his/her hand correctly for palmvein acquisition, the system starts to recognize the palmvein. When the system recognizes the user's identity, the word "Welcome!" appears, and the user's identity is displayed. Figure 12 shows that User 1 passes the authentication test.



Figure 12. The recognition result of legal user's (User 1's) palmvein.

When the user's palmvein cannot be recognized as a genuine user's palmvein, he/she cannot pass the authentication test, and the words "Sorry" and "No such person" appear, as shown in Figure 13.



Figure 13. The recognition result for the illegal user's palmvein.

To validate the performance of the implemented interactive system, this paper conducts some experiments using the completed system:

1. This paper conducts 30 experiments to validate the performance of the gesture recognition function. In 20 experiments, the users follow the system's instructions and make the correct gestures. In the other 10 experiments, users intentionally violate the system's instructions and make incorrect gestures. In a total of 30 experiments, the accuracy of the system's judgment on gesture recognition is 100%.

2. This paper also conducts 30 experiments to validate the performance of the hand tracking function. In 20 experiments, the users' hands intentionally leave the camera view range before palmvein recognition. In the other 10 experiments, the users' hands remain within the camera view range until the palmvein recognition has been completed. In the first 20 experiments, the corresponding time taken by the system for the user's hand to leave the camera view range is less than 1 s. In the last 10 experiments, the system successfully judges that the users' hands remain within the camera view range and do not leave.

The experimental results show that the interactive system implemented in this paper performs well and all functions can operate normally.

5. Conclusions and Future Work

This paper designs and implements an interactive system that achieves hand gesture recognition and hand tracking functions in an infrared environment. The application value of this system lies in the combination with palmvein recognition, which obviously improves the anti-counterfeiting performance during palmvein recognition. It can greatly reduce the possibility of forged palmveins, which can seriously threaten palmvein recognition systems. The hand gesture recognition implemented in this paper occurs against a pure background. In the future, we will improve the method by using a complex background. In addition, more different kinds of gestures will be collected to increase the diversity of the hand gestures.

Author Contributions: Each author discussed the details of the manuscript. J.X. designed and wrote the manuscript. J.X. implemented the proposed technique and provided the experimental results. L.L. provided some ideas and guidance. L.L. and B.-G.K. drafted and revised the manuscript. L.L. and B.-G.K. provided research funding support. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by the National Natural Science Foundation of China under Grant 61866028, Technology Innovation Guidance Program Project of Jiangxi Province (Special Project of Technology Cooperation), China under Grant 20212BDH81003.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due toprivacy issue.

Acknowledgments: We thank the anonymous reviewers for their valuable suggestions that improved the quality of this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Guo, L.; Lu, Z.; Yao, L. Human-machine interaction sensing technology based on hand gesture recognition: A review. *IEEE Trans. Hum.-Mach. Syst.* **2021**, *51*, 300–309. [CrossRef]
- Gomez-Barrero, M.; Drozdowski, P.; Rathgeb, C.; Patino, J.; Todisco, M.; Nautsch, A.; Damer, N.; Priesnitz, J.; Evans, N.; Busch, C. Biometrics in the Era of COVID-19: Challenges and Opportunities. *IEEE Trans. Technol. Soc.* 2022, *3*, 307–322. [CrossRef]
- Masi, I.; Wu, Y.; Hassner, T.; Natarajan, P. Deep face recognition: A survey. In Proceedings of the 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, Brazil, 29 October–1 November 2018; pp. 471–478.
- 4. Jia, W.; Zhang, B.; Lu, J.; Zhu, Y.; Zhao, Y.; Zuo, W.; Ling, H. Palmprint recognition based on complete direction representation. *IEEE Trans. Image Process.* 2017, 26, 4483–4498. [CrossRef] [PubMed]
- 5. Fei, L.; Lu, G.; Jia, W.; Teng, S.; Zhang, D. Feature extraction methods for palmprint recognition: A survey and evaluation. *IEEE Trans. Syst. Man Cybern. Syst.* 2018, 49, 346–363. [CrossRef]
- Zhong, D.; Du, X.; Zhong, K. Decade progress of palmprint recognition: A brief survey. *Neurocomputing* 2019, 328, 16–28. [CrossRef]
- Leng, L.; Teoh, A.B.J. Alignment-free row-co-occurrence cancelable palmprint fuzzy vault. *Pattern Recognit.* 2015, 48, 2290–2303. [CrossRef]

- 8. Qin, H.; El-Yacoubi, M.A.; Li, Y.; Liu, C. Multi-scale and multi-direction GAN for CNN-based single palm-vein identification. *IEEE Trans. Inf. Forensics Secur.* 2021, *16*, 2652–2666. [CrossRef]
- Magadia, A.P.I.D.; Zamora, R.F.G.L.; Linsangan, N.B.; Angelia, H.L.P. Bimodal hand vein recognition system using support vector machine. In Proceedings of the 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Manila, Philippines, 3–7 December 2020; pp. 1–5.
- Wu, T.; Leng, L.; Khan, M.K.; Khan, F.A. Palmprint-palmvein fusion recognition based on deep hashing network. *IEEE Access* 2021, 9, 135816–135827. [CrossRef]
- 11. Qin, H.; Gong, C.; Li, Y.; Gao, X.; El-Yacoubi, M.A. Label enhancement-based multiscale transformer for palm-vein recognition. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–17. [CrossRef]
- 12. Sandhya, T.; Reddy, G.S. An optimized elman neural network for contactless palm-vein recognition framework. *Wirel. Pers. Commun.* 2023, 131, 2773–2795. [CrossRef]
- 13. Sun, Y.; Wang, C. Presentation attacks in palmprint recognition systems. J. Multimed. Inf. Syst. 2022, 9, 103–112. [CrossRef]
- 14. Wang, F.; Leng, L.; Teoh, A.B.J.; Chu, J. Palmprint false acceptance attack with a generative adversarial network (GAN). *Appl. Sci.* **2020**, *10*, 8547. [CrossRef]
- 15. Erden, F.; Cetin, A.E. Hand gesture based remote control system using infrared sensors and a camera. *IEEE Trans. Consum. Electron.* **2014**, *60*, 675–680. [CrossRef]
- 16. Yu, L.; Abuella, H.; Islam, M.Z.; O'Hara, J.F.; Crick, C.; Ekin, S. Gesture recognition using reflected visible and infrared lightwave signals. *IEEE Trans. Hum. Mach. Syst.* 2021, *51*, 44–55. [CrossRef]
- García-Bautista, G.; Trujillo-Romero, F.; Caballero-Morales, S.O. Mexican sign language recognition using kinect and data time warping algorithm. In Proceedings of the 2017 International Conference on Electronics, Communications and Computers (CONIELECOMP), Cholula, Mexico, 22–24 February 2017; pp. 1–5.
- Mantecón, T.; Del-Blanco, C.R.; Jaureguizar, F.; García, N. A real-time gesture recognition system using near-infrared imagery. PLoS ONE 2019, 14, e0223320. [CrossRef] [PubMed]
- Kumar, P.; Gauba, H.; Roy, P.P.; Dogra, D.P. A multimodal framework for sensor based sign language recognition. *Neurocomputing* 2017, 259, 21–38. [CrossRef]
- Xu, H.; Leng, L.; Yang, Z.; Teoh, A.B.J.; Jin, Z. Multi-task pre-training with soft biometrics for transfer-learning palmprint recognition. *Neural Process. Lett.* 2023, 55, 2341–2358. [CrossRef]
- 21. Park, H.J.; Kang, J.W.; Kim, B.G. ssFPN: Scale sequence (S 2) feature-based feature pyramid network for object detection. *Sensors* 2023, 23, 4432. [CrossRef]
- 22. Mujahid, A.; Awan, M.J.; Yasin, A.; Mohammed, M.A.; Damaševičius, R.; Maskeliūnas, R.; Abdulkareem, K.H. Real-time hand gesture recognition based on deep learning YOLOv3 model. *Appl. Sci.* **2021**, *11*, 4164. [CrossRef]
- 23. Qi, W.; Ovur, S.; Li, Z.; Marzullo, A.; Song, R. Multi-sensor guided hand gesture recognition for a teleoperated robot using a recurrent neural network. *IEEE Robot. Autom. Lett.* **2021**, *6*, 6039–6045. [CrossRef]
- 24. Jangpangi, M.; Kumar, S.; Bhardwaj, D.; Kim, B.G.; Roy, P.P. Handwriting recognition using wasserstein metric in adversarial learning. *SN Comput. Sci.* 2022, *4*, 43. [CrossRef]
- Yang, Z.; Leng, L.; Wu, T.; Li, M.; Chu, J. Multi-order texture features for palmprint recognition. Artif. Intell. Rev. 2023, 56, 995–1011. [CrossRef]
- Leng, L.; Li, M.; Kim, C.; Bi, X. Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. *Multimed. Tools Appl.* 2017, 76, 333–354. [CrossRef]
- 27. Leng, L.; Zhang, J. Palmhash code vs. palmphasor code. *Neurocomputing* **2013**, *108*, 1–12. [CrossRef]
- 28. Sahoo, J.P.; Prakash, A.J.; Pławiak, P.; Samantray, S. Real-time hand gesture recognition using fine-tuned convolutional neural network. *Sensors* 2022, 22, 706. [CrossRef] [PubMed]
- Saboo, S.; Singha, J. Vision based two-level hand tracking system for dynamic hand gestures in indoor environment. *Multimed. Tools Appl.* 2021, 80, 20579–20598. [CrossRef]
- Kulshreshth, A.; Zorn, C.; LaViola, J.J. Poster: Real-time markerless kinect based finger tracking and hand gesture recognition for HCI. In Proceedings of the 2013 IEEE Symposium on 3D User Interfaces (3DUI), Orlando, FL, USA, 16–17 March 2013; pp. 187–188.
- 31. Houston, A.; Walters, V.; Corbett, T.; Coppack, R. Evaluation of a multi-sensor Leap Motion setup for biomechanical motion capture of the hand. *J. Biomech.* **2021**, *127*, 110713. [CrossRef] [PubMed]
- 32. Ovur, S.E.; Su, H.; Qi, W.; De Momi, E.; Ferrigno, G. Novel adaptive sensor fusion methodology for hand pose estimation with multileap motion. *IEEE Trans. Instrum. Meas.* 2021, 70, 1–8. [CrossRef]
- 33. Nonnarit, O.; Ratchatanantakit, N.; Tangnimitchok, S.; Ortega, F.; Barreto, A.; et al. Hand tracking interface for virtual reality interaction based on marg sensors. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; pp. 1717–1722.
- Santoni, F.; De Angelis, A.; Moschitta, A.; Carbone, P. MagIK: A hand-tracking magnetic positioning system based on a kinematic model of the hand. *IEEE Trans. Instrum. Meas.* 2021, 70, 1–13. [CrossRef]
- Mueller, F.; Bernard, F.; Sotnychenko, O.; Mehta, D.; Sridhar, S.; Casas, D.; Theobalt, C. Ganerated hands for real-time 3d hand tracking from monocular rgb. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 49–59.

- 36. Han, S.; Liu, B.; Cabezas, R.; Twigg, C.D.; Zhang, P.; Petkau, J.; Yu, T.H.; Tai, C.J.; Akbay, M.; Wang, Z.; et al. MEgATrack: Monochrome egocentric articulated hand-tracking for virtual reality. *ACM Trans. Graph. (ToG)* **2020**, *39*, 87:1–87:13. [CrossRef]
- Menghani, G. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. ACM Comput. Surv. 2023, 55, 1–37. [CrossRef]
- Sharifani, K.; Amini, M. Machine learning and deep learning: A review of methods and applications. World Inf. Technol. Eng. J. 2023, 10, 3897–3904.
- Zhuang, Q.; Gan, S.; Zhang, L. Human-computer interaction based health diagnostics using ResNet34 for tongue image classification. *Comput. Methods Programs Biomed.* 2022, 226, 107096. [CrossRef] [PubMed]
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18-22 June 2018; pp. 7132–7141.
- 42. Huang, J.; Ren, L.; Zhou, X.; Yan, K. An improved neural network based on SENet for sleep stage classification. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 4948–4956. [CrossRef] [PubMed]
- Zhang, F.; Bazarevsky, V.; Vakunov, A.; Tkachenka, A.; Sung, G.; Chang, C.L.; Grundmann, M. Mediapipe hands: On-device real-time hand tracking. *arXiv* 2020, arXiv: 2006.10214.
- Ghanbari, S.; Ashtyani, Z.P.; Masouleh, M.T. User identification based on hand geometrical biometrics using media-pipe. In Proceedings of the 2022 30th International Conference on Electrical Engineering (ICEE), Seoul, Korea, 17–19 May 2022; pp. 373–378.
- Güney, G.; Jansen, T.S.; Dill, S.; Schulz, J.B.; Dafotakis, M.; Hoog Antink, C.; Braczynski, A.K. Video-based hand movement analysis of parkinson patients before and after medication using high-frame-rate videos and MediaPipe. *Sensors* 2022, 22, 7992. [CrossRef] [PubMed]
- Peiming, G.; Shiwei, L.; Liyin, S.; Xiyu, H.; Zhiyuan, Z.; Mingzhe, C.; Zhenzhen, L. A PyQt5-based GUI for operational verification of wave forcasting system. In Proceedings of the 2020 International Conference on Information Science, Parallel and Distributed Systems (ISPDS), Xi'an, China, 14–16 August 2020; pp. 204–211.
- Renyi, L.; Yingzi, T. Semi-automatic marking system for robot competition based on PyQT5. In Proceedings of the 2021 International Conference on Intelligent Computing, Automation and Systems (ICICAS), Chongqing, China, 29–30 December 2021; pp. 251–254.
- 48. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv: 1409.1556.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- 50. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A convnet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 11976–11986.
- Chen, J.; Kao, S.h.; He, H.; Zhuo, W.; Wen, S.; Lee, C.H.; Chan, S.H.G. Run, don't walk: Chasing higher FLOPS for faster neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 12021–12031.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.