The Journal of Supercomputing (2023) 79:2767–2782
https://doi.org/10.1007/s11227-022-04680-4

Anti-aliasing convolution neural network of finger vein recognition for virtual reality (VR) human–robot equipment of metaverse

Nghi C. Tran[1] · Jian-Hong Wang[2] · Toan H. Vu[1] · Tzu-Chiang Tai[3] · Jia-Ching Wang[1]

Accepted: 20 June 2022 / Published online: 22 August 2022

Abstract

Metaverse, which is anticipated to be the future of the internet, is a 3D virtual world in which users interact via highly customizable computer avatars. It is considerably promising for several industries, including gaming, education, and business. However, it still has drawbacks, particularly in the privacy and identity threads. When a person joins the metaverse via a virtual reality (VR) human-robot equipment, their avatar, digital assets, and private information may be compromised by cybercriminals. This paper introduces a specific Finger Vein Recognition approach for the virtual reality (VR) human-robot equipment of the metaverse of the Metaverse to prevent others from misappropriating it. Finger vein is a is a biometric feature hidden beneath our skin. It is considerably more secure in person verification than other hand-based biometric characteristics such as finger print and palm print since it is difficult to imitate. Most conventional finger vein recognition systems that use hand-crafted features are ineffective, especially for images with low quality, low contrast, scale variation, translation, and rotation. Deep learning methods have been demonstrated to be more successful than traditional methods in computer vision. This paper develops a finger vein recognition system based on a convolution neural network and anti-aliasing technique. We employ/ utilize a contrast image enhancement algorithm in the preprocessing step to improve performance of the system. The proposed approach is evaluated on three publicly available finger vein datasets. Experimental results show that our proposed method outperforms the current state-of-the-art methods, improvement of 97.66% accuracy on FVUSM dataset, 99.94% accuracy on SDUMLA dataset, and 88.19% accuracy on THUFV2 dataset.

Keywords Metaverse · Virtual reality (VR) human–robot · Biometrics · Finger vein recognition · Deep learning · Convolution network · Pre-processing · Image processing · Anti-aliasing

Extended author information available on the last page of the article

1 Introduction

Metaverse is a 3D virtual world of online social networking and a persistent and decentralized 3D virtual environment. It relates to a collection of technological devices that are related to the Internet of Things (IoT), blockchain, artificial intelligence (AI), and many other areas. Data in a metaverse has a special identification tag and can be used as traceable data in a blockchain-based system system. It provides a useful resource for artificial intelligence [1]. Users may access the metaverse virtual space through virtual reality glasses, augmented reality glasses, mobile phones, personal computers, and video gaming consoles. Despite demonstrating a promising future in the business, education, retail, and real estate industries, the metaverse is restricted by the technical limitations of hardware devices, the necessity for sensors to interact with a real-time virtual world, information privacy problems, and user addiction.

Dionisio, J.D. (2013) [2]discuss the transition from a collection of isolated virtual worlds to an integrated network of 3D virtual worlds, or a metaverse. Four important factors are considered to define success in attempt advancement of a fully functioning metaverse: realism (the user's realistic psychological and emotional experience in the metaverse), ubiquity (the accessibility of current digital devices and the preservation of the user's virtual identity), interoperability (the synchronicity and continuity of digital assets and user movement), and scalability (the system efficiently when a large number of users are active/online in the metaverse). Beside the difficulties of its concept and technology, the development of the metaverse may be hampered by substantial economic and political restrictions [2].

In addition to theoretical and evolutionary study of metaverse, different investigation also focus on practical applications based on the metaverse platform. For examples, due to the demands of Covid 19's disease prevention criteria, more people prefer the disease consultations to be conducted virtually rather than in person. Han et al. [3] develop a realistic virtual area for diagnosis and therapy based on the metaverse platform. This research concentrates on strengthening the patient's comfort throughout online diagnosis and therapy. The Physically Based Rendering (PBR) is used for simulating a virtual environment that resembles the real world. It contributes to enhancing the user experience in the virtual world.

Second Life, a free 3D virtual world where users can interact and communicate in real-time, serves as the initial version of the metaverse. However, a novel concept of the metaverse is necessarily suggested due to the innovation of social networks. Park et al. [4] identifies three areas where a new metaverse differs from the previous one: hardware (headset device, GPU memory, network), software (development of the identification and expression model that takes advantage of the hardware's parallelism), and contents (the availability of content that people and participate in Metaverse). The authors also offer a thorough analysis of the software and technologies that might provide social meaning in a metaverse in three different approaches: user interaction, implementation, and application.

Lee et al. [5] introduce a brain-to-speech (BTS) system for smart communication utilizing brain signals in the real world. The research reveals a potential future use of brain-computer interfaces in the Metaverse system: speech-based smart home control via discussion with a virtual assistant.

The metaverse demonstrates that it is a virtual environment with promising contributions in which people will interact, share information, and participate in activities with one another using digital avatars. However, this new form of communication raises a number of privacy and security problems, particularly cyber threats, fraud, and identity theft, which poses a threat to the digital world. As a result, developing a convenient and reliable identity verification system to protect our 3D gadgets and identities in the metaverse from identity theft is challenging. One way to protect users from fraud or device theft is to utilize biometric identification technologies. Biometrics utilizes a registered person's unique traits, such as fingerprints or iris scans, to verify the real identity of a login attempt and grant access to a user's account. Kuo Wang and Ajay Kumar [6] use iris recognition to verify user identification in metaverse. The author addresses significant constraints of existing technologies for off-angle iris identification from closer distances utilizing head-mounted devices and obtaining state-of-the-art iris recognition. Similar to the iris, every individual's finger veins are unique. Because of the internal nature of the finger vein, it is difficult to copy or forge. Veins exist only in live humans and their patterns are unique and do not change throughout life. The vein patterns can be captured using illumination of near infrared light with no touching required. Compared to traditional authentication systems such as password or swipe card, biometric systems (e.g., face recognition [7], iris recognition [8], finger vein recognition [9], etc.) are considered to be more secure. This study suggests a specific finger vein recognition for the virtual reality (VR) human-robot equipment of the metaverse in an effort to prevent misappropriation of the VR human-robot technology.

In conventional approaches, finger vein recognition systems include two stages: finger vein pattern extraction and identification. The former detects finger vein pattern from an input image, while the latter identifies vein identity. Repeated line tracking [9], maximum curvature [10], and wide line detector [11] are three early algorithms for extracting finger vein patterns. These algorithms detect vein patterns based on the cross sectional profile of the finger vein, where a vein appears darker than the surrounding region. The Repeated line tracking algorithm employs random initial positions and repeated tracking of vein lines in images, while the maximum curvature method determines vein pattern by locating local maximum curvatures in cross sectional profiles of finger veins. On the contrary, the Wide line detector technique can detect the vein line's width instead of detecting the vein line's center. After extracting the vein pattern, the finger vein image is identified by comparing the vein pattern with all the vein patterns in the database using a pattern matching algorithm such as structural matching [12] or template matching [13]. However, the aforementioned systems for finger vein recognition have certain drawbacks, such as

being difficult to build as an end-to-end system and image quality sensitive (images with low quality, low contrast, or variations in scaling, translation, and rotation).

Deep learning has demonstrated success in a variety of applications, including natural language processing [14], speech recognition [15], and object detection [16] . It also produced an impressive result in the field of biometric recognition [17], particularly on a finger veins recognition system. For example, the studies in [18] and [19] use Alex Net and VGG 16 architecture for finger vein recognition and achieve better results than the traditional algorithms. FV-GAN model, a Generative Adversarial Network (GAN) model for finger vein based on Cycle-GAN architecture, is introduced in [20] to overcome the low quality image problem and insufficient data. The authors in [21] suggests a fully convolutional neural network, an extension of U-Net, and an embedded conditional random field as an end-to-end system for pixel-wise finger vein segmentation, and achieves effective performance in vein pattern segmentation results. R.S. Kuze et al. [22] have recently proposed a novel CNN model that modifies Densenet-161 architecture by integrating a custom embedder module to the backbone model and achieves state-of-the-art results on publicly available finger vein datasets.

The studies on finger vein recognition based on deep learning algorithms showed promising results for the problem. However, the existing methods still do not cope well with the insufficient data. In this study, we develop a CNN-based finger vein recognition method using the pretrained model for addressing this problem. Our system integrates the blur pool layer into the pretrained model to improve the model performance, especially for data suffering from rotation or translations, which causes the shift variance problem. In addition, we apply dual exposure fusion algorithm [23], a preprocessing process that combines eye exposure adjustment and brain exposure adjustment, to enhance the image contrast and thus increase the recognition accuracy. Our proposed model is evaluated on three public finger vein datasets and the results are compared with the state of the art model.

The rest of the paper is organized as follows. Section 2 describes the proposed method for finger vein recognition. Section 3 introduce the hardware device. The experimental results are discussed in Sect. 4. Finally, the conclusions and remarks are given in Sect. 5.

2 The proposed method

2.1 Proposed system architecture

The finger vein recognition proposed in this study for determining vein identity includes three phases namely the training phase (Fig. 1a), the registration phase (Fig. 1b), and the recognition phase (Fig. 1c). First, in the training phase, a pretrained CNN model and a finger vein dataset are used to build a deep learning model, which is then used as a finger vein feature extractor. Then, the registration

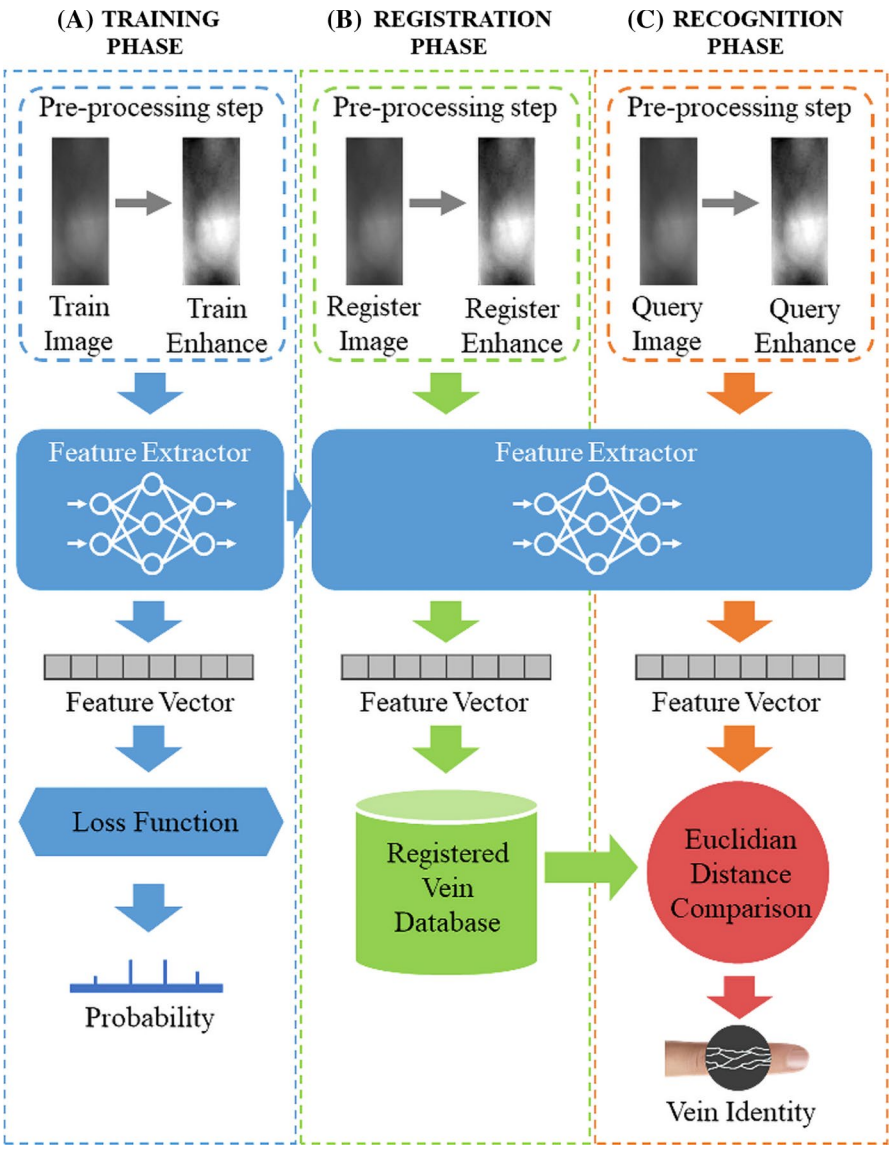

Fig. 1 Proposed finger vein recognition system **a** Train phase; **b** Registration phase; **c** Recognition phase

phase accepts the enrolled finger vein images and uses the deep learning model built in the training phase to extract the vein features. The extracted feature vectors are then stored in an enrolled finger vein database. Finally, the recognition phase aims to recognize the vein identity of the input finger vein image. The features of the input image are extracted and compared with the features of the registered finger veins based on the Euclidean distance.

This section focuses on preprocessing techniques, anti-aliasing convolution neural networks to extract the embedding vector, and loss function.

2.2 Preprocessing

In the acquisition step, finger vein images taken with infrared light are usually sensitive to ambient illumination. These images can be underexposed or overexposed, resulting in poor recognition performance. Therefore, it is necessary to apply image processing techniques for enhancing image quality to improve recognition performance. We propose to use a dual exposure fusion algorithm [23] to enhance the contrast of input finger vein images before using these images for training, registration, and recognition tasks. This algorithm has been shown to be a powerful technique for enhancing images in low light conditions, which may help to achieve better results in recognition tasks. The contrast enhancement process in this algorithm involves two stages. The eye exposure adjustment stage generates a multi-exposure image set, and the brain exposure adjustment stage fuses the generated image set into the final enhancement image. Figure 2 shows the resulting image after applying this contrast enhancement algorithm.

2.3 Anti-aliasing convolutional neural network

Convolutional neural networks extract output feature maps from input images for the classification tasks. Shifting features' position in input images can result in different output feature maps. Pooling layers provide an approach to down sampling feature maps to reduce the sensitivity of features' position. However, the variance problem still remains due to the down sampling process. For example, a pooling layer with stride 2 can correctly extract even pixel shifting while it is incorrect for odd pixel shifting. Thus, if shifting the input affects the output feature map, the CNN model

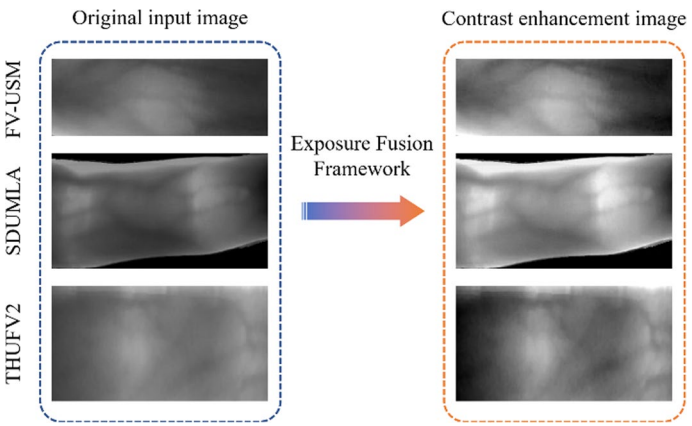

Fig. 2 Image contrast enhancement using Exposure Fusion Framework

is a shift variance. Otherwise, the CNN model is shift invariance. Equations 1 and 2 show the cases of shift variance and shift invariance, respectively [24].

$$Shift_{\Delta h,\Delta w}\left(\tilde{F}(X)\right) = \tilde{F}\left(Shift_{\Delta h,\Delta w}(X)\right)\forall(\Delta h, \Delta w) \tag{1}$$

$$\tilde{F}(X) = \tilde{F}\left(Shift_{\Delta h,\Delta w}(X)\right)\forall(\Delta h, \Delta w) \tag{2}$$

where $X \in \mathbb{R}^{H\times W\times 3}$ represents input image. Theeature map extracted by CNN model is $F(X) \in \mathbb{R}^{H_i\times W_i\times C_i}$, with spatial resolution $H_i \times W_i$ and C_i channels. The original resoon $\tilde{F}(X) \in \mathbb{R}^{H\times W\times C_i}$ is up-sampled by the feature map. $\Delta h, \Delta w$ are shifting distance.

To improve the shift equivariance, we use the blur pool layer in a convolution neural network to make it shift-invariance [24]. Blur pool layers replace the original pooling layers to blur the input feature before the down sampling step, and thereby minimize the influence of the shift of the input feature. The technique of blurring an image in order to smooth out the detail is also known as "anti-aliasing," thus we refer to it as an anti-aliasing convolutional neural network. The proposed model is constructed in two main steps. In the first step, the Max Pool layer (stride 2) in the original model is replaced by the Max Blur Pool block, which consists of the Max Pool layer (stride 1) and the Blur Pool layer (stride 2) (Eq. 3).

$$\begin{aligned} &MaxPool_{stride=2} \rightarrow MaxBlurPool_{stride=2} \\ &= MaxPool_{stride=1} \circ BlurPool_{stride=2} \end{aligned} \tag{3}$$

In the second step, the Blur Pool layer (stride 2) replaces the Average Pool layer (stride 2) in original model (Eq. 4).

$$AvgPool_{stride=2} \rightarrow BlurPool_{stride=2} \tag{4}$$

2.4 Loss function

In our system, the CNN model is used to extract embedding vectors from input images for the classification tasks. The model needs to minimize the intra-class distance as well as maximize the interclass distance. For this reason, the Additive Angular Margin Penalty (AAMP) [25] is used instead of the soft max loss because it gives better results in feature extraction. The formula is shown in Eq. 5.

$$L = -\frac{1}{B}\sum_{i=1}^{B}\log\frac{e^{s\cos\left(\theta_{y_i}+m\right)}}{e^{s\cos\left(\theta_{y_i}+m\right)} + \sum_{j=1,j\neq y_i}^{U} e^{s\cos\theta_j}} \tag{5}$$

where $x_i \in \mathbb{R}^d$ is deep feature of the i-th sample that belongs to the y_i-th class. The d is dimension embedding feature. $W_j \in \mathbb{R}^d$ represent the j-th column of the weight $W \in \mathbb{R}^d$ and $b_j \in \mathbb{R}^d$ represent the bias term. B and U are the batch size and class number, respectively. The scaling factor and the penalty margin are represented by hyper parameters s and m. B is batch size (the value is 8), U is the class number (it

depends on the dataset, for example, its value is 305 for *THUFV2 dataset*, 318 for *SDUMLA dataset*, and 246 for FV-*USM dataset*), the s value is 30.0, and the m value is 0.50.

3 Hardware device

There are several ways to enter the metaverse, including computers and mobile phones. However, these gadgets do not produce the same level of interactivity as realistic metaverse devices. Virtual reality equipment (Google, Samsung, HTC Vive, etc.) is the most immersive method. In this part, we describe the via a virtual reality (VR) human-robot equipment (HTC Vive) for accessing the metaverse, which enables interaction with a digital environment. Furthermore, we introduce the finger vein verification device for finger vein identification.

3.1 Virtual reality (VR) human–robot equipment of metaverse

HTC Vive is one of the earliest devices that allows users to thoroughly immerse themselves in the metaverse. It consists of a headset and a controller. Figure 3 depicts the design structure of the HTC Vive device, with Fig. 3a depicting the headgear and Fig. 3b, c and d depicting the controller. Figure 3b is left-hand-side trackpad that moves the player character, touch and hold the trackpad until the character has reached the desired location. Figure 3c is right/left trigger that press the left/right trigger to select objects or options. Figure 3d is right grip button that press the right grip button or right trigger to grab the components of a generator. Table 1 is the HTC Vive headset specification. The HTC Vive headgear gives a VR field of vision that is affected by the player's motions, and the triggers and trackpads on the controllers are used to hold objects/select locations and control the player character's movements, respectively.

3.2 The verification device of finger vein identification

In the finger vein identification system, finger vein images are acquired using a specific device, a finger vein scanner. Because veins are not visible under normal lighting conditions, this device captures them using Near-Infrared (NIR) light. The image of the finger vein is then passed through a feature extractor module to produce the appropriate feature vector, which is subsequently stored in the database of enrolled finger veins. When the user login to the system, an image of his/her finger vein is extracted and compared to the registered finger veins in order to validate the correct user. The advantages of finger vein authentication are as follows:

- Biometric technology has the ability to not lose, not to be stolen, and to have no burden of remembering passwords.

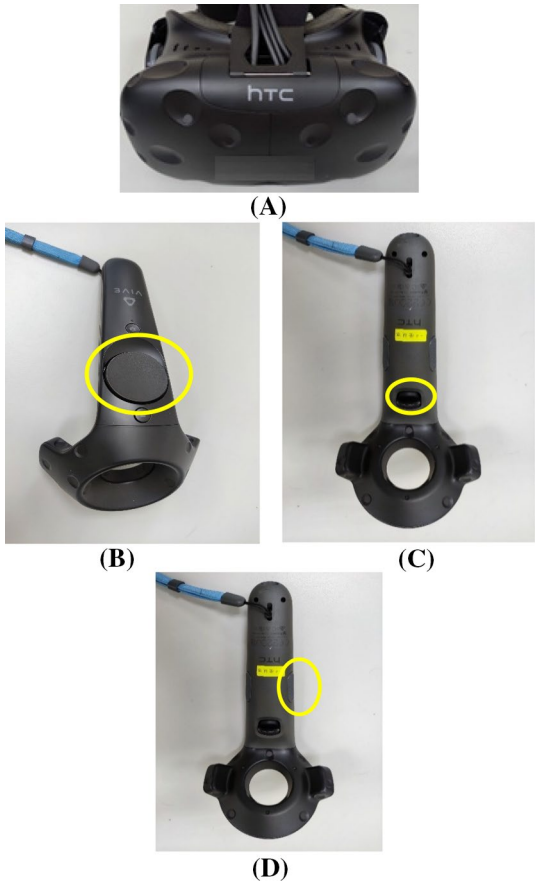

Fig. 3 HTC Vive headset and controllers. **a** Headset; **b** Left-hand-side trackpad; **c** Right/left trigger; **d** Right grip button

Table 1 The HTC Vive headset specification

Field of view	110°
Resolution	2160×1200 pixels (both eyes)
Tracing technology	Outside-in tracking
Space requirements	A minimum of 2 m×1.5 m Recommended play area: 3.5 m×3.5 m
Screen refresh rate	90 Hz
Supported degrees of freedom	The headset and controllers support 6 degrees of freedom

OS requirements: Windows™ 7 SP1, Windows™ 8.1, or Windows™ 10

GPU: NVIDIA GeForceTM GTX 1060, AMD RadeonTM RX 480 equivalent or better

CPU: Intel™ Core™ i5-4590 or AMD FX™ 8350 equivalent or better

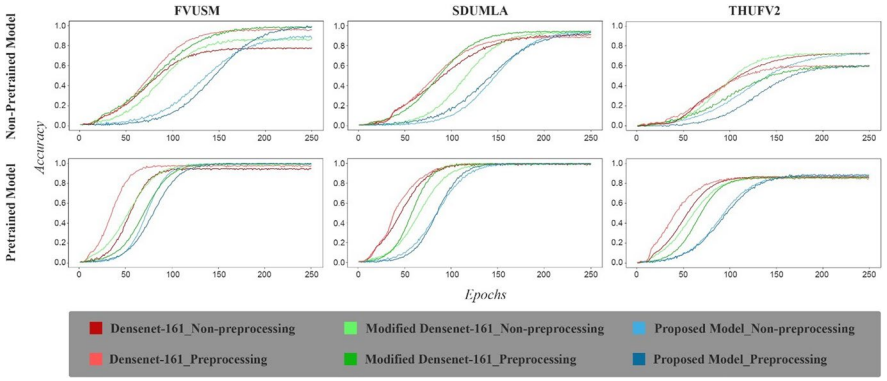

Fig. 4 The accuracy of finger vein recognition system in different datasets

- The finger veins belong to the internal information of the human body and are not affected by the roughness of the epidermis and the external environment (temperature, humidity).
- Finger vein authentication has high accuracy, cannot be copied, cannot be forged, and is safe and convenient.
- Non-living fingers cannot obtain vein image features, and cannot identify and forge.
- Finger veins are used for identity authentication. When acquiring finger vein images, the identification can be completed without touching the finger with the device.

4 Experiments

4.1 Datasets

To evaluate our approach, we employed three publicly available datasets on finger veins: FV-USM [26], THU FVFDT2 [27], and SDUMLA [28] . Their properties are explained in more detail in the section below.

- *The* FV-*USM dataset* [26] *("Data Availability Statement")*: This dataset was published by Sains University Malaysia. It was collected from a total of 123 people's index and middle fingers on their left and right hands, including 83 men and 40 women aged from 20 to 52 years. Images were taken in two sessions, with six images taken per finger in each session. It provided the extracted ROI images with a resolution of 100×300 pixels for finger vein recognition.
- *The SDUMLA dataset* [28] *("Data Availability Statement")*: There are 106 participants were enrolled in the Shan-Dong University finger vein dataset (SDUMLA). Each participant contributed index, middle, and ring fingers from two hands, resulting in a total of 636 finger classes. Images of the captured fin-

gers are collected in a single session. Each finger provided 6 images during the session, resulting in a total of 3816 images. The spatial resolution of the original images is 320×240 pixels. Because their work does not provide the ROI finger vein image, preprocessed the dataset to obtain ROI images with a size of 300×150 pixels.

- *The THU* FVFDT2 *dataset* [27] *("Data Availability Statement")*: Tsinghua University published the THUFV2 dataset in 2014. It included Regions of Interest (ROIs) of finger vein and dorsal textures from 610 different subjects, which corresponded to 610 images of each type. Images were captured in two separate sessions. Each ROI was normalized to 200×100 pixels. Most of the subjects were Tsinghua University Graduate School students and staff volunteers from Shenzhen.

Our research team ensures that all research data are available. For each dataset, the training and validation set and the test set are split into two equal parts. In the FV-USM and THU-FV2 datasets, the first part is used for training and validation set, while the second part is used as test set. In the SDUMLA dataset, the first half of the labels is used as the training and validation sets, whereas the second half of the labels is used as the test set. The training and validation set is then divided into training set and validation set for use in training phase with an 8:2 ratio. Table 2 describes the detail of finger vein datasets.

4.2 Experimental configuration

Our model is developed based on Densenet-161 architecture [29]. It reuses weights from pretrained ImageNet to shorten training time and overcome the insufficient training data. Then, a custom embedder module is added to model for extracting features vectors. Finally, blur pool layers introduced in Sect. 2.2 are integrated to make model shift invariance. The proposed model architecture is represented in Table 3.

In the training phase, each dataset is divided into two different sets using the open set scenario: training and validation set, and test set as explained in Sect. 4.1. Model parameters are adjusted by using stochastic gradient descent (SGD) and one cycle learning rate scheduler during the training stage in 250 epochs. To ensure a reliable comparison, the same configuration is used for all evaluation models.

Table 2 Details of the finger vein datasets

Dataset	# of classes	# of Sessions	# of images per class	Total samples
FV-USM	492	2	12 images (6 images per session)	5904
SDUMLA	636	1	6 images	3816
THU-FV2	610	2	2 images (1 image per session)	1220

Table 3 Anti-aliasing modified densenet-161 architecture

Layers	
Convolution	7×7 conv, stride 2
Max BlurPool	3×3 max pool, stride 1
	3×3 max pool, stride 1
DenseBlock 1	[1×1 conv, 3×3 conv]×6
Transition 1	1×1 conv
	4×4 blur pool, stride 2
DenseBlock 2	[1×1 conv, 3×3 conv]×12
Transition 2	1×1 conv
	4×4 blur pool, stride 2
DenseBlock 3	[1×1 conv, 3×3 conv]×36
Transition 3	1×1 conv
	4×4 blur pool, stride 2
DenseBlock 4	[1×1 conv, 3×3 conv]×24
Custom embedder module	7×7 global avg pool
	batch-normalization
	Dropout
	full-connected (feature embedder)
	batch-normalization
Classification layer	full-connected (classifier)

4.3 Experimental results

To evaluate the proposed system, EER (Equal Error Rate) and accuracy are used as evaluation metrics. EER metric is the most popular evaluation metric used in biometric authentication system evaluation [30]. It is the point at which the false acceptance rate and the false rejection rate are equal. The lower EER value is the better system's performance. Accuracy metric measures identification performance of biometric system. It refers to a one to many matching situation. A queried finger vein is compared with all registered veins in the registered database to detect the correct vein identity. The number of queried vein images and enrolled vein images varies depending on the dataset. Each class in FV-USM dataset has 6 registered images and 6 queried images, whereas SDUMLA dataset has 3 registered images and 3 queried images per class, and THU-FV2 dataset has just 1 registered image and 1 queried image per class.

Table 4 shows the experimental results of Densenet-161 [29], modified Densenet-161 [22] , and proposed model with preprocessing and without preprocessing. Two tests were taken on each model on three public finger vein datasets: the first test was conducted without pretrained weights and the second test was conducted with pretrained weights. In the first test, i.e., with non-pretrained weights, the performance of proposed model is considerably higher than Densenet-161 and modified Densenet-161 model in FV-USM dataset and SDUMLA dataset but slightly worse than the Densenet-161 in THUFV2 dataset

Table 4 The experimental results of Densenet-161, modified Densenet-161 [11], and proposed model with preprocessing and without preprocessing

Pretrained	Preprocessing	Model	FVUSM		SDUMLA		THUFV2	
			Accuracy (%)	EER (%)	Accuracy (%)	EER (%)	Accuracy (%)	EER (%)
No	No	Densenet-161	75.07	8.23	91.09	2.47	72.13	19.44
		Modified Densenet-161	83.74	5.45	93.61	1.97	71.79	20.93
		Proposed model	86.99	4.85	95.07	1.93	72.06	19.74
	Yes	Densenet-161	92.95	3.18	88.68	3.31	60.00	17.59
		Modified Densenet-161	95.80	2.83	94.34	1.99	59.67	17.75
		Proposed model	96.88	2.68	94.03	2.48	60.03	17.58
Yes	No	Densenet-161	92.55	3.45	99.37	0.55	86.15	15.10
		Modified Densenet-161	96.88	2.25	99.49	0.41	85.25	16.45
		Proposed model	97.28	2.18	99.89	0.31	87.37	14.93
	Yes	Densenet-161	95.59	2.57	99.79	0.44	84.91	16.84
		Modified Densenet-161	97.42	2.16	99.75	0.38	85.56	16.05
		Proposed model	**97.66**	**2.03**	**99.94**	**0.24**	**88.19**	**12.61**

with preprocessing and non-preprocessing. In the second test, i.e., with pertained weights, the performance of proposed model outperforms Densenet-161 and modified Densenet-161 model in all datasets with preprocessing and non-preprocessing. The preprocessing technique illustrates that it can improve the model's performance, especially for non-pretrained weights with nearly 10% accuracy. However, in THUFV2 dataset with non-pretrained weights, it is not effective because of the low number of samples (1 sample per class and 1220 samples totally) and insufficient training epochs. Moreover, Figure 4 demonstrates the effectiveness of pretrained model in enhancing model performance and decreasing training time compare to non-pretrained model. Our proposed system outperforms the state-of-the-art model in all experiments. For datasets with insufficient data (THUFV2 dataset), combining pre-trained weight and preprocessing approach significantly improves model accuracy (72.06% to 88.19%). However, the augmentation strategy can be used to increase the model's performance and make it more robust.

5 Conclusion

This study proposes a special Finger Vein Recognition for the via a virtual reality (VR) human-robot equipment of the Metaverse in order to prevent others from misappropriating the via a virtual reality (VR) human-robot equipment of the metaverse. We propose a deep learning strategy for finger vein recognition that is based on an anti-aliasing convolution neural network and dual exposure fusion algorithm. The model is trained using the pretrained weight on the ImageNet dataset with the same training protocol and hyperparameters. The experimental results show that our system outperforms the existing state-of-the-art model on the three different public datasets. To achieve the improvement, we conducted experiments using various model structures and preprocessing techniques. The proposed models, which incorporate a blur pool layer and the dual exposure fusion algorithm, have never been employed previously in the finger vein recognition system. Thus, its implementation for vein recognition is novel. As a result, our proposed model increased both reliability and popularity in finger vein recognition.

Acknowledgements This study was supported by the scientific research funds of Shandong University of Technology, Zibo, China. Jian-Hong Wang received his PhD degree in Communications Engineering from the National Chung Cheng University, Taiwan, in January 2015. He is currently a professor at School of Computer Science and Technology, Shandong University of Technology, Zibo, China. Jian-Hong Wang is the corresponding author of this paper and can be contacted at: wwwccucomtw@gmail.com or jhwang_2015@163.com.

Data availability statement The datasets generated during and analyzed during the current study are available in the FV-*USM dataset* [26] repository, [http://drfendi.com/fv_usm_database/]. The datasets generated during and analyzed during the current study are available in the *SDUMLA dataset* [28] repository, [https://tsapps.nist.gov/BDbC/Search/Details/420]. The datasets generated during and analyzed

during the current study are available in the *THUFV2 dataset* [27] repository, [https://www.sigs.tsinghua.edu.cn/labs/vipl/thu-fvfdt.html].

Declarations

Conflict of interest The authors declare no conflicts of interest regarding the publication of this paper.

6. References

1. Mozumder MAI, Sheeraz MM, Athar A, Aich S, Kim HC (2022) Overview: technology roadmap of the future trend of metaverse based on IoT, Blockchain, AI Technique, and Medical Domain Metaverse Activity. In: 2022 24th International Conference on Advanced Communication Technology (ICACT), 2022, pp 256–261, doi: https://doi.org/10.23919/ICACT53585.2022.9728808.
2. Dionisio JD, Burns WG, Gilbert R (2013) 3D virtual worlds and the metaverse: current status and future possibilities. ACM Comput Surv 45:341–3438
3. Han Y, Oh S (2021) Investigation and research on the negotiation space of mental and mental illness based on metaverse. In: 2021 International Conference on Information and Communication Technology Convergence (ICTC), 2021, pp 673–677. doi:https://doi.org/10.1109/ICTC52510.2021.9621118
4. Park SM, Kim YG (2022) A metaverse: taxonomy, components, applications, and open challenges. IEEE Access 10:4209–4251. https://doi.org/10.1109/ACCESS.2021.3140175
5. Lee SH, Lee YE, Lee SW (2022) Toward imagined speech based smart communication system: potential applications on metaverse conditions. In: 2022 10th International Winter Conference on Brain Computer Interface (BCI), pp 1–4, doi:https://doi.org/10.1109/BCI53720.2022.9734827.
6. Wang K, Kumar A (2022) Human identification in metaverse using egocentric iris recognition
7. Tayal Y, Kumar Pandey P, Singh DBV (2013) Face recognition using eigenface. Int J Emerg Technol Comput Appl Sci 3(1):50–55
8. Daugman J (2009) How iris recognition works. In: The essential guide to image processing, pp 715–739. Elsevier
9. Miura N, Nagasaka A, Miyatake T (2004) Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. Mach Vis Appl 15(4):194–203
10. Miura N, Nagasaka A, Miyatake T (2007) Extraction of finger vein patterns using maximum curvature points in image profiles. IEICE Trans Inf Syst 90(8):1185–1194
11. Huang B, Dai Y, Li R, Tang D, Li W (2010) Finger vein authentication based on wide line detector and pattern normalization. In: 2010 20th International Conference on Pattern Recognition. IEEE, 2010, pp 1269–1272
12. Zhang W, Wang Y (2002) "Core based structure matching algorithm of Finger Vein verification," in Object recognition supported by user interaction for service robots. IEEE 1:70–74
13. Nagao M (1983) Methods of image pattern recognition. Corona, San Antonio, TX
14. Otter DW, Medina JR, Kalita JK (2020) A survey of the usages of deep learning for natural language processing. IEEE Trans Ne ural Netw Learn Syst 32(2):604–624
15. Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K (2019) Speech recognition using deep neural networks: a systematic review. I EEE Access 7:19143–19165
16. Wu X, Sahoo D, Hoi SCH (2020) Recent advances in deep learning for object detection. Neurocomputing 396:39–64
17. Minaee S, Abdolrashidi A, Su H, Bennamoun M, Zhang D (2019) Biometrics recognition using deep learning: a survey. arXiv:1912.00271
18. Liu W, Li W, Sun L, Zhang L, Chen P (2017) Finger vein recognition based on deep learning. In: 2017 12th IEEE Conference o n Industrial Electronics and Applications (ICIEA). IEEE, 2017, pp 205–210
19. Hong HG, Lee MB, Park KR (2017) "Convolutional neural network based finger vein recognition using nir image sensors". Sensors 17(6):1297
20. Yang W, Hui C, Chen Z, Xue J-H, Liao Q (2019) Fvgan: finger vein representation using generative adversarial networks. IEEE Trans Inf Forens Secur 14(9):2512–2524
21. Zeng J et al., (2020) "Finger vein verification algorithm based on fully convolutional neural network and conditional random field". IEEE Access 8:65402–65419

22. Rıdvan Salih Kuzu, Emanuele Maiboranay, and Patrizio Campisi, "Vein-based biometric verification using transfer learning," in 2020 43rd International Conferenceon Telecommunications and Signal Processing (TSP). IEEE, 2020, pp. 403–409.
23. Ying Z, Li G, Gao W (2017) A bioinspired multi-exposure fusion framework for low-light image enhancement. arxiv:1711.00591
24. Zhang R (2019) Making convolutional networks shift invariant again. In: International Conference on Machine Learning. PML R, pp 7324–7334
25. Deng J, Guo J, Xue N, Zafeiriou S (2019) Arcface: additive angular margin loss for deep face recognition. In: Proceedings of th e IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 4690–4699
26. Asaari MSM, Suandi SA, Rosdi BA (2014) Fusion of band limited phase only correlation and width centroid contour distance fo r finger based biometrics. Exp Syst Appl 41(7):3367–3382
27. Yang W, Qin C, Liao Q (2014) A database with ROI extraction for studying fusion of finger vein and finger dorsal texture. In: C hinese Conference on Biometric Recognition. Springer, pp 266–270
28. Yin Y, Liu L, Sun X (2011) SDUMLAHMT: a multimodal biometric database. In: Chinese Conference on Biometric Recognitio n. Springer, pp 260–268
29. Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger (2017) Densely Connected Convolutional Networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
30. Jain AK, Flynn P, Ross AA (2007) Handbook of biometrics. Springer Science & Business Media

Authors and Affiliations

Nghi C. Tran[1] · Jian-Hong Wang[2] · Toan H. Vu[1] · Tzu-Chiang Tai[3] · Jia-Ching Wang[1]

✉ Jian-Hong Wang
wwwccucomtw@gmail.com

[1] Department of Computer Science and Information Engineering, National Central University, Taoyuan City, Taiwan

[2] School of Computer Science and Technology, Shandong University of Technology, Zibo 255049, Shandong, China

[3] Department of Computer Science and Information Engineering, Providence University, Taichung City, Taiwan