MYU Tokyo

S & M 4149

# Biometrics Security: Lightweight Finger-vein Recognition Based on Efficient Focal Aggregation Block and Vision Transformer

Hua-Ching Chen,[1] Liang-Ying Ke,[2] and Chih-Hsien Hsia[3,4*]

[1]School of Information Engineering, Xiamen Ocean Vocational College, Fujian 361100, China
[2]Department of Engineering Science, National Cheng Kung University, Tainan 701, Taiwan
[3]Department of Computer Science and Information Engineering, National Ilan University, Yilan 26007, Taiwan
[4]Department of Business Administration, Chaoyang University of Technology, Taichung 413310, Taiwan

(Received November 18, 2024; accepted August 14, 2025)

***Keywords:*** deep learning, finger-vein recognition, vision transformer, lightweight network, smart home

The rapid development of IoT, cloud computing, and AI in recent years has benefited smart homes tremendously. However, camera footage showing facial images of users from smart homes has raised security hazards. When a user's facial image is stolen, it can undermine the security of facial data verification. An effective alternative solution is to replace biometrics based on facial data with those based on finger-vein features. Finger-vein biometrics are difficult to counterfeit, steal, or wear out. However, current technology for recognizing finger-vein characteristics is limited by challenges in extracting features when using a fewer number of parameters, which tends to decrease the model's recognition performance. To address these problems, we propose a lightweight efficient focal aggregation model for finger-vein recognition (EFA-FV), which is based on the efficient focal aggregation block (EFAB) and vision transformer (ViT). The EFAB module not only lets the EFA-FV model effectively extract global features from finger-vein characteristics through the ViT architecture, but it also provides the proposed model with the generalization capability characteristic of a convolutional neural network model. As a result, the EFA-FV model with fewer parameters can be smoothly trained on a database with relatively few samples, enhancing the performance of the finger-vein recognition model. The experimental results indicate that the proposed finger-vein model achieved correct identification rates of 99.90 and 99.83% on the FV-USM and MMCBNU-6000 public databases, respectively, while maintaining a smaller number of parameters of only about 0.60 M. This makes it the most successful system available in comparison with those in previous studies.

## 1. Introduction

Recent developments in the IoT, cloud computing, and AI have allowed the technology in the smart home domain to advance extensively, prompting significant attention from researchers.(1) Smart home systems can monitor family members' activities through numerous cameras

*Corresponding author: e-mail: hsiach@niu.edu.tw
https://doi.org/10.18494/SAM5483

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

positioned around the home, allowing users to keep track of household activities at all times. These cameras can recognize users' hand gestures to facilitate interaction with various devices.[2,3] However, the deployment of many cameras in smart homes for monitoring users' daily behaviors exposes users' facial images to risks of theft, potentially compromising the security of identity recognition systems based on facial data being collected.[4] Furthermore, the physiological or behavioral traits used in biometric recognition technology possess characteristics such as universality, distinctiveness, permanence, and collectability, which make users' biometric features difficult to modify once they are stolen. This leads to a prolonged risk of identity theft, thus raising concerns about the privacy and security of biometric recognition technologies based on extrinsic biological characteristics.[5] In contrast, finger-vein features located beneath the surface of the skin are difficult to steal, forge, or wear down, which makes finger-vein recognition an effective alternative to traditional identity verification methods.[6] However, vein-based biometric recognition technologies still have some drawbacks, mainly from the methods used to capture finger-vein images. Low-cost near-infrared (NIR) cameras are typically used to acquire finger-vein images, but poor lighting conditions could lead to overexposure or insufficient brightness, resulting in missing finger-vein features. Additionally, if users rotate or move their fingers while finger-vein images are being taken, the amount of infrared exposure can change, affecting the captured features. These issues result in varied finger-vein features, leading to errors in identity recognition.

To address the issues of image translation and illumination variations, which lead to the loss of vein features, and to develop a lightweight model suitable for practical finger-vein recognition applications, in this study, we developed a lightweight efficient focal aggregation model for finger-vein recognition (EFA-FV), which is constructed on the basis of the efficient focal aggregation block (EFAB) and vision transformer (ViT) architectures. The proposed finger-vein recognition model not only effectively extracts finger-vein features but also significantly reduces model parameters, thereby enhancing the feasibility of deploying the model in real-world environments. The contributions of this study are summarized as follows:

- In this study, we developed a lightweight deep-learning model for finger-vein recognition named EFA-FV, which is based on the EFAB and ViT architectures. The proposed model effectively extracts global finger-vein features while maintaining a smaller number of model parameters, thereby stabilizing the recognition performance of the finger-vein model.
- An efficient focal aggregation module (EFAM) with a multiscale feature aggregation mechanism is introduced to effectively capture the long-range dependency (LRD) of finger-vein images, thereby enhancing the adaptability of the model to image variations such as illumination changes and finger rotations.
- The proposed finger-vein recognition model achieves correct identification rates (*CIR*s) of 99.90 and 99.83% on the FV-USM and MMCBNU-6000 databases, respectively, while effectively reducing the number of model parameters to approximately 0.60 M. Consequently, the model is particularly suitable for deployment in resource-constrained yet security-critical applications, such as smart homes and mobile devices.

## 2. Related Work

To enhance the feature extraction capability of the proposed EFA-FV model, in this study, we investigated and analyzed previous research in finger-vein recognition. Hsia *et al.*[7] proposed an improved lightweight convolution neural network (ILCNN) for finger-vein recognition, which has a smaller number of model parameters. Although the improvement provided significant advantages for deployment in real-world environments, it increased the risk of losing feature information, thereby limiting the model's capability for feature extraction. On the other hand, Liu *et al.*[8] introduced the multiscale and multistage residual attention network (MMRAN) for finger-vein recognition. The MMRAN model enhanced the diversity of vein features through a residual architecture, but its complex design extended the gradient path, resulting in reduced learning efficiency.[9] Devkota and Kim[10] proposed an SE-DenseNet-HP model that combined a channel attention mechanism and hybrid pooling for finger-vein recognition. Since this method employed a DenseNet model with high computational complexity[11] for extracting finger-vein features, it is prone to overfitting issues when there is insufficient training data. Furthermore, Ke and Hsia[12] introduced a lightweight dual-attention convolutional neural network (LDA-FV) constructed through a dual-attention-based inverted residual block. Although the LDA-FV model enhances the extraction capability of global features through an attention mechanism, it still cannot handle LRD information; thus, it is difficult for the LDA-FV model to cope with complex environmental changes in images. Li and Zhang[13] proposed an FV-ViT model based on the ViT architecture, which captures the multiscale features of finger veins and the LRD information between features through the multihead self-attention mechanism (MHSA).[14] Through this, the model's capability to extract the structure and features of finger veins is enhanced. However, the use of the MHSA within the vanilla ViT for feature extraction significantly increased the model's computational complexity, leading to difficulties in deploying the model in practical environments.

Previous finger-vein models for designing lightweight architectures have primarily employed multibranch structures to enhance the model's capability to capture diverse finger vein features or utilized MHSA mechanisms to capture the LRD information of finger vein features. These approaches aim to mitigate feature loss caused by translation or illumination variations in finger-vein images. However, increasing the number of model branches or incorporating MHSA mechanisms often leads to higher computational complexity and a larger number of parameters, thereby increasing the costs associated with model training and deployment. To address these challenges, in this study, we developed an EFAM module capable of multiscale feature aggregation and global feature perception. This module effectively enhances the model's capability to learn diverse finger vein features while simultaneously capturing LRD information, thereby improving the stability of feature recognition in finger vein models.

## 3. Proposed Lightweight Finger-vein Recognition Model

### 3.1 Overall architecture

To enable the finger-vein recognition model to maintain a small number of parameters while effectively extracting features from finger-vein images, we developed the EFA-FV model in this study. It was constructed on the basis of the EFAB module and the ViT architecture for finger-vein recognition. The modules within each stage of the EFA-FV model were stacked at a 1:1:3:1 ratio to effectively enhance the *CIR* of the model.(15) Our proposed EFA-FV model is shown in Fig. 1(a).

Initially, the model converts finger-vein images into patches using overlapping patch embeddings constructed using depthwise separable convolutions.(16) This transformation enables the EFA-FV model to exploit the inherent inductive bias properties of the convolutional architecture. The local correlation of the finger-vein feature map is then enhanced, as well as the model's generalization capability. Subsequently, the EFA-FV model extracts global features from the finger-vein feature map through the proposed EFAB module, allowing the model to effectively capture important features in the finger-vein image. This results in a robust and effective finger-vein model. Figure 1(b) demonstrates the EFAB module within the EFA-FV model.

The EFAB module first normalizes the finger-vein features and then utilizes the EFAM to aggregate multiscale finger-vein features. Through this, the original finger-vein features are modulated accordingly. This process enables the EFA-FV model to filter important finger-vein

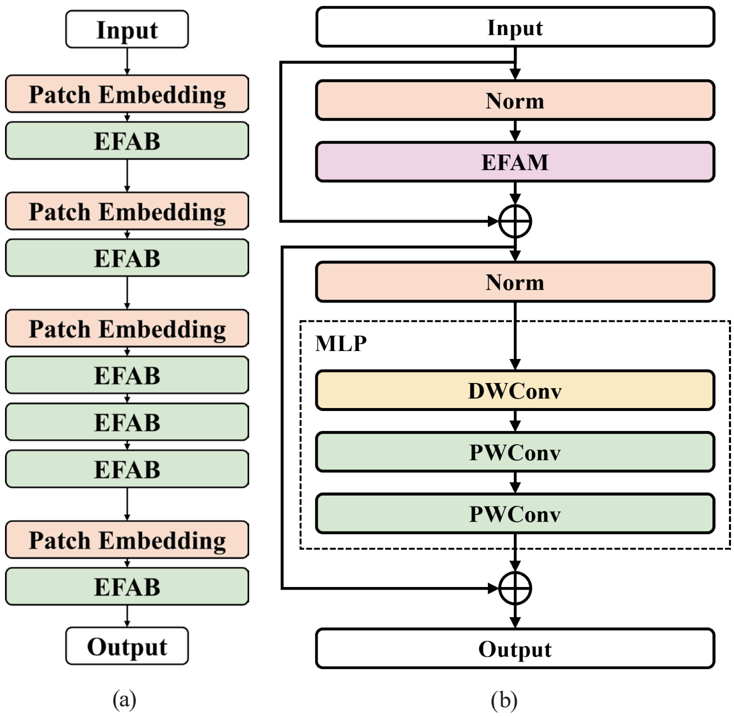


Fig. 1. (Color online) Finger-vein recognition model developed in this study. (a) Diagram of the EFA-FV model. (b) Diagram of the EFAB module within the EFA-FV model.

features using multiscale information. These operations allow the EFA-FV model to have the capability of the ViT architecture to capture LRD information and to retain the generalization capability of the convolutional neural network (CNN) architecture.[17] Then, the EFAB module normalizes the extracted global features of the finger-vein features by employing a multilayer perceptron (MLP) to map the features into a high-dimensional space for feature extraction. This approach not only enhances the feature representation capability of the EFA-FV model but also effectively captures the relationships among the finger-vein features.

### 3.2 EFAM

To enable the proposed EFA-FV model to extract global features while maintaining an efficient training process, in this paper, we introduce an EFAM module capable of aggregating multiscale features and effectively perceiving LRD information, as shown in Fig. 2. This module integrates the inductive bias of CNN with the capability of ViT to extract global features, enabling the EFA-FV model to effectively learn the correlations among finger-vein features and enhance its feature representation capability for finger-vein recognition. Initially, the EFAM module performs feature transformation on the finger-vein features using pointwise convolution (PWConv)[16] to map the features into a higher-dimensional feature space while enabling effective feature extraction. Subsequently, the EFAM module divides the transformed finger-vein features on the basis of channel dimension. Then, these are used as query features and the basis for feature aggregation in subsequent computations. During the feature aggregation phase, the EFAM module extracts spatial dimension information from the finger-vein features using

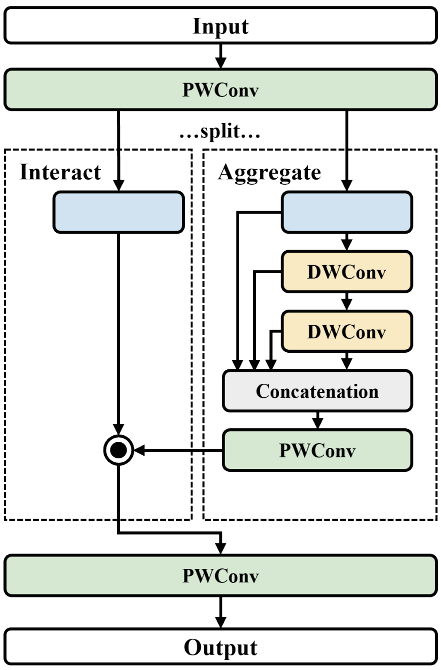


Fig. 2. (Color online) Proposed architecture of EFAM.

depthwise convolution (DWConv).[16] This enables the model to learn more discriminative spatial features, thereby enhancing its representation capability. By employing DWConv with different kernel sizes, the EFAM module can extract multiscale finger-vein features and perceive wide-range features from the finger-vein feature map; thus, the model's understanding of finger-vein images is further improved. Next, the EFAM module merges the finger-vein features of different scales and inputs the combined features into PWConv for feature transformation. This process integrates the extracted finger-vein features from different kernel sizes and produces an attention map for the subsequent filtering of important finger-vein features. In the feature interaction phase, the EFAM module utilizes the Hadamard product to combine the attention map generated during the feature aggregation phase with the query features. This allows the selection of important finger-vein features. Finally, the EFAM module applies PWConv to the filtered finger-vein features for feature transformation, enhancing the model's feature representation capability and improving the recognition performance of the EFA-FV model.

## 4. Experiment Results and Comparison

### 4.1 Finger-vein databases

The current study utilized the FV-USM[18,19] and MMCBNU-6000 public databases[20] for training and testing the proposed EFA-FV model. The following sections provide a detailed introduction to these public databases.

#### 4.1.1 FV-USM Database[18]

For this database, a NIR LED with a wavelength of 850 nm was employed to collect finger-vein images from 123 participants ($N = 83$ for males and $N = 40$ for females) aged 20 to 52 years. The fingers used for collecting the infrared finger-vein images were the 1) left index, 2) left middle, 3) right index, and 4) right middle fingers. These fingers are considered distinct classes, resulting in a total of 492 classes within the database. Six images of each finger were collected in a session, and each participant attended another session two weeks later. The database has a total of 5904 infrared finger-vein images with a resolution of 640 × 480. Before training the model, the database was divided into training, validation, and test sets for model evaluation at a ratio of 4:1:1 for each dataset.

#### 4.1.2 MMCBNU-6000 Database[20]

Similar to the above, this database used an NIR LED with a wavelength of 850 nm to collect infrared finger-vein images with a resolution of 640 × 480. Samples were collected from 100 participants ($N = 83$ for males and $N = 17$ for females) aged 16 to 72 years from 20 different countries. For the database, 10 finger-vein images were collected for each participant's left and right index, middle, and ring fingers. These fingers were considered distinct classes, resulting in 600 identity classes and 6000 finger-vein images. Before training the model, the database was divided into training, validation, and test sets for model evaluation, with a ratio of 3:1:1 for each dataset.

### 4.2 Experimental setting and results

In terms of hyperparameters, in this study, we set the image size to 112, the batch size to 32, and the number of epochs to 40, using AdamW as the model optimizer, with an initial learning rate set to 0.0002. Finally, the overall finger-vein model was trained using the PyTorch Toolbox on a system equipped with an NVIDIA RTX 3090 Ti graphics processing unit. Furthermore, to assess the overall performance of the model, *CIR* was used as the metric to evaluate the model's security, as shown in Eq. (1). A higher *CIR* indicates greater security in recognition, whereas a lower *CIR* suggests reduced security.

$$CIR \quad \frac{\text{Number of correctly identified cases}}{\text{Total number of identified cases}} \times 100 \tag{1}$$

The feasibility of the proposed EFA-FV model in the finger-vein recognition task was validated using the FV-USM public database. The experimental results showed that the EFA-FV model achieved a *CIR* of 99.90% on the FV-USM database, with a total of only 0.60 million parameters, as shown in Table 1. Compared with previous state-of-the-art finger-vein models, the EFA-FV model not only effectively reduced the number of parameters but also maintained a high *CIR*. This means that the EFA-FV model can perform identity recognition effectively while providing better computational efficiency, making it more suitable for deployment in practical application scenarios. Furthermore, the results revealed that the EFA-FV model can effectively extract discriminative features from low-quality finger-vein images and use them for accurate identity recognition, demonstrating its superior performance in handling low-quality images.

The identification accuracy of the EFA-FV across a diverse age distribution in the database was evaluated using the MMCBNU-6000 database. From the experimental results, the proposed EFA-FV model achieved a *CIR* of 99.83% on the MMCBNU-6000 database, as shown in Table 2. Compared with previous state-of-the-art finger-vein models, the EFA-FV model proposed in this paper had an improvement of 0.13% in terms of the *CIR* metric. On the basis of the above results, it can be concluded that the proposed EFA-FV model can effectively perform identity

Table 1
Comparison of the proposed model with those in previous research on the FV-USM database. In the table, *CIR* denotes the correct identification rate, and Params denote the number of parameters.

| Method | *CIR* (%) | Params (M) |
|---|---|---|
| ILCNN[7] | 99.82 | 1.23 |
| MMRAN[8] | 96.07 | 3.51 |
| LDA-FV[12] | 99.90 | 1.20 |
| Semi-PFVN[21] | 94.67 | 3.35 |
| EfficientNet-B0[22] | 99.70 | 4.64 |
| FV-RSA[23] | 99.90 | 8.70 |
| EFA-FV (This work) | 99.90 | 0.60 |

Table 2
Comparison of the proposed model with those in previous research on the MMCBNU-6000 database.

| Method | *CIR* (%) |
|---|---|
| ILCNN[7] | 99.25 |
| LDA-FV[12] | 99.70 |
| EfficientNet-B0[22] | 89.44 |
| AGCNN[24] | 91.06 |
| PVTv2-B0[25] | 86.53 |
| EFA-FV (This work) | 99.83 |

Table 3
Ablation studies on the FV-USM database to evaluate the effectiveness of the proposed EFAM. Specifically, the experimental conditions included training finger-vein recognition models with and without EFAM, denoted as w/ EFAM (using the proposed EFAM) and w/o EFAM (using the original MHSA mechanism), respectively.

| Method | EFAM | *CIR* (%) | Params (M) |
|---|---|---|---|
| EFA-FV(This work) | w/o | 99.70 | 0.84 |
| | w/ | 99.90 | 0.60 |

Table 4
Ablation studies on the MMCBNU-6000 database to evaluate the effectiveness of the proposed EFAM.

| Methods | EFAM | *CIR* (%) | Params (M) |
|---|---|---|---|
| EFA-FV(This work) | w/o | 99.58 | 0.86 |
| | w/ | 99.83 | 0.61 |

recognition on a database with a wide age distribution, further revealing its recognition capability for users across different age groups.

To verify the effectiveness of the proposed EFAM, in this study, we designed two experimental conditions on the basis of a traditional MHSA mechanism and the EFAM and performed ablation studies using two public databases: FV-USM and MMCBNU-6000. Experimental results demonstrated that the finger-vein recognition model incorporating the proposed EFAM achieved superior recognition performance, achieving *CIR*s of 99.90 and 99.83% on the FV-USM and MMCBNU-6000 databases, respectively, as shown in Tables 3 and 4. Specifically, compared with the traditional MHSA mechanism, the proposed EFAM improved the *CIR* by 0.20% on FV-USM and 0.25% on MMCBNU-6000, while simultaneously reducing the number of parameters by 0.24 and 0.25 M, respectively. These results highlight the fact that the proposed EFAM not only enhances the feature representation capability of the finger-vein recognition model by leveraging the inductive bias inherent in CNN architectures but also effectively achieves model lightweighting.

## 5. Conclusions

In this study, we developed an EFA-FV architecture on the basis of the EFAB module and the ViT architecture for lightweight finger-vein recognition. The performance of the EFA-FV model was evaluated using both the FV-USM and MMCBNU-6000 public databases simultaneously. The EFAB module proposed in this paper enabled the EFA-FV model to obtain the characteristics of the ViT architecture, extract global features of finger-vein characteristics, and achieve the generalization capability of a CNN. Owing to this, the EFA-FV model was effectively trained on databases with limited samples while simultaneously enhancing the performance of the finger-vein recognition model. According to the experimental results, the proposed EFA-FV model achieved *CIR*s of 99.90% on the FV-USM database and 99.83% on the MMCBNU-6000 database, with only about 0.60 M parameters. These results indicate that the EFA-FV model not only demonstrated superior *CIR*s on both the FV-USM and MMCBNU-6000 public databases but also maintained a small number of parameters. Owing to the lightweight design of the EFA-FV model and its high *CIR*s, it is less susceptible to hardware limitations and is well-suited for high-security application scenarios, such as smart homes and mobile devices.

## References

1 J. Yu, A. d. Antonio, and E. Villalba-Mora: Computers **11** (2022) 26. https://doi.org/10.3390/computers11020026
2 A. Rahim, Y. Zhong, T. Ahmad, S. Ahmad, P. Pławiak, and M. Hammad: Sensors **23** (2023) 6979. https://doi.org/10.3390/s23156979
3 B. I. Alabdullah, H. Ansar, N. A. Mudawi, A. Alazeb, A. Alshahrani, S. S. Alotaibi, and A. Jalal: Sensors **23** (2023) 7523. https://doi.org/10.3390/s23177523
4 D. Kagan, G. F. Alpert, and M. Fire: IEEE Trans. Comput. Social Syst. **11** (2024) 933. https://doi.org/10.1109/TCSS.2022.3231987
5 E. Wenger, S. Shan, H. Zheng, and B. Y. Zhao: IEEE Symp. Security and Privacy (IEEE, 2023) 864–881.
6 B. Ma, K. Wang, and Y. Hu: Sci. Rep. **13** (2023) 249. https://doi.org/10.1038/s41598-023-27524-4
7 C.-H. Hsia, L.-Y. Ke, and S.-T. Chen: Bioengineering **10** (2023) 919. https://doi.org/10.3390/bioengineering10080919
8 W. Liu, H. Lu, Y. Wang, Y. Li, Z. Qu, and Y. Li: Appl. Intell. **53** (2023) 3273. https://doi.org/10.1007/s10489-022-03645-7
9 C.-Y. Wang, H.-Y. M. Liao, and I.-H. Yeh: arXiv preprint arXiv:2211.04800 (arXiv, 2022). https://doi.org/10.48550/arXiv.2211.04800
10 N. Devkota and B. W. Kim: Electronics **13** (2024) 501. https://doi.org/10.3390/electronics13030501
11 G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger: IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2017) 4700–4708.
12 L.-Y. Ke and C.-H. Hsia: Sens. Mater. **36** (2024) 945. https://doi.org/10.18494/SAM4858
13 X. Li and B.-B. Zhang: IEEE Access **11** (2023) 75451. https://doi.org/10.1109/ACCESS.2023.3297212
14 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin: Neural Information Processing Systems (IEEE, 2017) 5998–6008.
15 Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie: IEEE/CVF Conf. Computer Vision and Pattern Recognition (IEEE, 2022) 11976–11986.
16 A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam: arXiv preprint arXiv:1704.04861 (arXiv, 2017). https://doi.org/10.48550/arXiv.1704.04861
17 Z. Dai, H. Liu, Q. V. Le, and M. Tan: Neural Information Processing Systems (IEEE, 2021) 3965–3977.
18 M. S. M. Asaari, S. A. Suandi, and B. A. Rosdi: Expert Syst. Appl. **41** (2014) 3367. https://doi.org/10.1016/j.eswa.2013.11.033
19 C.-H. Hsia and C.-F. Lai: Sens. Mater. **32** (2020) 3221. https://doi.org/10.18494/SAM.2020.2861
20 Y. Lu, S. J. Xie, S. Yoon, Z. Wang, and D. S. Park: Int. Congr. Image and Signal Processing (IEEE, 2013) 410–415.
21 T. Chai, J. Li, S. Prasad, Q. Lu, and Z. Zhang: J. Inf. Secur. Appl. **67** (2022) 103211. https://doi.org/10.1016/j.jisa.2022.103211
22 M. Tan and Q. V. Le: Int. Conf. Machine Learning (ICML, 2019). https://doi.org/10.48550/arXiv.1905.11946
23 Z. Zhang, G. Chen, W. Zhang, and H. Wang: IEEE Access **12** (2024) 1943. https://doi.org/10.1109/ACCESS.2023.3347922
24 Y. Zhang, W. Li, L. Zhang, X. Ning, L. Sun, and Y. Lu: Concurrency Comput. Pract. Exp. **34** (2022) e5697. https://doi.org/10.1002/cpe.5697
25 W. Wang, E. Xie, X. Li, D.-P. Fan, K. Song, D. Liang, T. Lu, P. Luo, and L. Shao: Comp. Visual Media **8** (2022) 415. https://doi.org/10.1007/s41095-022-0274-8

## About the Authors

**Hua-Ching Chen** received his B.S. and M.S. degrees in Business Administration and Computer Science and Information Engineering from National Quemoy University, Kinmen, Taiwan, R.O.C., in 2008 and 2010, respectively. He received his Ph.D. degree in Electronic Engineering from Xiamen University, Xiamen, Fujian, China, P.R.C., in 2014. He is currently a lecturer at the Department of Computer Science and Information Engineering, National Quemoy University. His current research interests include wireless networks, cloud computing, fuzzy systems, optimal learning algorithms, image processing, and robot systems.

**Liang-Ying Ke** received his B.S. and M.S. degrees in Computer Science and Information Engineering from National Ilan University, Yilan, Taiwan, in 2022 and 2024, respectively. He is currently pursuing his Ph.D. degree in Engineering Science at National Cheng Kung University, Tainan, Taiwan. His research interests include computer vision, deep learning, biometrics recognition, and virtual beauty technologies.

**Chih-Hsien Hsia** received his first Ph.D. degree in Electrical and Computer Engineering from Tamkang University, New Taipei City, Taiwan, and his second Ph.D. degree from National Cheng Kung University, Tainan City, Taiwan, in 2010 and 2023, respectively. In 2007, he was a visiting scholar at Iowa State University, Ames, IA, USA. From 2010 to 2013, he was a postdoctoral research fellow at the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan. From 2013 to 2015, he was an assistant professor at the Department of Electrical Engineering at Chinese Culture University (CCU), Taipei City, Taiwan. He was an associate professor at CCU and National Ilan University (NIU) from 2015 to 2017. From 2019 to 2020, he was the director of the Research Planning Division, Research & Development, NIU, Yilan County, Taiwan. He is currently a distinguished professor at the Department of Computer Science and Information Engineering, NIU, the Director of the Multimedia & Intelligent Technical Lab., the CEO of the AI Promotion Office, the Director of the AIoX Research Center, NIU, and an honorary distinguished professor at Chaoyang University of Technology, Taichung City, Taiwan. His research interests include AI/GAI in Multimedia, DSP IC Design, and Cognitive Engineering. Dr. Hsia received multiple honors, including a Fellow of the International Association of Advanced Materials (IAAM) in 2025, and the Outstanding Young Scholar Award from the Taiwan Association of Systems Science and Engineering (TASSE) in 2020 and from the Computer Society of the Republic of China (CSROC) in 2018. He is the Chapter Chair of the IEEE Young Professionals Group and the Vice Chair of the IEEE Signal Processing Society, Taipei Section, and the Vice President of the IET Taipei Local Network. He has served as an associate editor of the Journal of Imaging Science and Technology, Scientific Reports, the Journal of Computers, and the Journal of Intelligent Communication.