IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 January 2024, accepted 17 March 2024, date of publication 27 March 2024, date of current version 2 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3382197

RESEARCH ARTICLE

Lite-HDNet: A Lightweight Domain-Adaptive Segmentation Framework for Improved Finger Vein Pattern Extraction

YINGXIN LI[1], YUCONG CHEN[1], JUNYING ZENG[2], (Member, IEEE), CHUANBO QIN[2], AND WENGUANG ZHANG[3]

[1]Department of Network Information, Jiangmen Central Hospital, Jiangmen, Guangdong 529020, China
[2]Department of Intelligent Manufacturing, Wuyi University, Jiangmen, Guangdong 529020, China
[3]Department of Neurosurgery, Jiangmen Central Hospital, Jiangmen, Guangdong 529020, China

Corresponding author: Wenguang Zhang (zhangwenguang@jmszxyy.com.cn)

This work was supported in part by the Special Project in Key Areas of Artificial Intelligence in Guangdong Universities under Grant 2019KZDZX1017, in part by the 2022 Guangdong Provincial Education Department Graduate Education Innovation Project through Guangdong Education and Research Letter under Grant 2022 (No. 1), in part by the 2022 Guangdong Province Joint Training Graduate Demonstration Base Project YJS-SFJD-22-01, and in part by the 2023 Guangdong Provincial Medical Research Fund Project B2023100.

ABSTRACT Recent times have witnessed significant progress in deep learning-based finger vein pattern extraction methods, but two unavoidable issues still remain to be addressed. One is that the model trained on a single finger vein dataset shows poor generalizability, and the model performance is limited by the image quality of the single dataset; the other is that it is hard for the deep model to extract real-time finger vein patterns because of its large number of parameters and poor real-time performance. To address the aforementioned issues, we propose a novel lightweight domain-adaptive segmentation framework (Lite-HDNet) that learns a generic representation of different domains to improve the extraction of finger vein patterns. We propose a multi-domain feature knowledge transfer strategy and a domain migration loss converter to enable the trunk network to learn the robust representations of different finger vein datasets as well as to compensate for the heterogeneity between them. In the proposed framework, two lightweight segmentation networks are designed as the trunk branch and the auxiliary branch to achieve real-time extraction of finger vein patterns. Our approach has been extensively tested on four finger vein datasets available to the public, and the results show that our Lite-HDNet not only improves segmentation performance on all datasets but also effectively reduces heterogeneity between different domains. In addition, we also validated the real-time performance of Lite-HDNet on NVIDIA embedded terminals, proving the outperformance of our approach compared with previous lightweight segmentation networks.

INDEX TERMS Image segmentation, domain adaptation, finger vein extraction, knowledge transfer.

I. INTRODUCTION

With the increasing demand for security systems on the market, more and more attention has been paid to biometric identification, which is becoming one of the most critical and challenging tasks in information security. Because of its live recognition, high uniqueness, and strong anti-counterfeit properties, finger vein recognition technology has been widely used in information security, network payment, and other fields. Finger vein recognition technology mainly includes two key steps: feature extraction and matching recognition. The former is the basis for the latter, and extracting a clearer finger vein pattern can effectively improve the accuracy of matching recognition. Therefore, how to extract clear finger vein patterns has become a hot problem that many researchers are concerned with and devoted to solving.

Previous methods for extracting finger vein patterns mainly divide vein regions and background regions by some

The associate editor coordinating the review of this manuscript and approving it for publication was Shovan Barma.

attribute assumptions (such as valleys and straight lines), including repeated line tracking (RLT) (Miura et al. [1]), regional growth (Qin et al. [2]), maximum curvature point (MC) (Miura et al. [3]), Gabor filters (Yang et al. [4]), and wide line detectors (WLD) (Huang et al. [5]), etc. These traditional methods have some drawbacks: (1) Most methods perform poorly on low-quality images; (2) Finger vein patterns are marked by different thresholds, and it is difficult to determine these different thresholds, which makes it difficult to distinguish vein regions from non-vein regions; (3) Some methods to extract the finger vein pattern may exist with a vein line break. Now, many researchers have employed deep learning-based image segmentation techniques for the extraction of finger vein patterns in order to avoid the shortcomings of conventional approaches. Because the features of the vein are extracted directly from the original image in the deep learning method, the crucial error of model extraction is minimized. In addition, the rich prior knowledge can be fully utilized by forming a large number of digital vein images, which also resolves the problem of class imbalance.

Current deep learning-based approaches for extracting finger veins have two key issues: weak model generalizability and subpar real-time performance. Many segmentation networks can achieve good results on a single finger vein dataset through training but perform poorly on other datasets or practical applications because of the poor generalizability of the model due to a single training dataset, which limits the segmentation performance of the model. An effective way to improve the generalizability of the model is to include multiple finger vein datasets in the training phase of the network, which can not only enrich the diversity of finger vein features but also improve the segmentation performance of the model. However, due to the heterogeneity between different finger vein datasets, direct training between finger vein images from different domains is undesirable, and direct use of mixed datasets to train the model does not yield good results. As a result, one of the topics of our research is how to effectively reduce heterogeneity between finger vein images in order to improve the model's generalizability. In practical applications, finger vein recognition systems need to have a low manufacturing cost, a good effect, a fast response, etc. Finger vein feature extraction is one of the key steps in finger vein recognition technology, and the real-time performance of finger vein extraction will directly affect the real-time performance of the whole recognition process, while many deep learning-based finger vein extraction methods have low real-time performance and are difficult to deploy on low-cost embedded platforms. Therefore, another focus of our research is to extract finger vein patterns in real time on embedded platforms.

To solve the above problems, we propose a novel lightweight domain-adaptive segmentation framework (Lite-HDNet) that utilizes knowledge distillation to transfer feature information from multiple finger vein datasets to a trunk network, and instructs the trunk network to train so that the trunk network can learn a generic representation of finger veins, address the heterogeneity between different finger vein datasets, and improve the generalizability and real-time performance of the model. For learning a more robust representation in the model, we propose a middle-layer guided multi-source feature knowledge transfer strategy (MDFKT) so that the trunk network can learn feature knowledge from different domains in a focused manner. It is not an easy task to solve the problem of heterogeneity between different domains of data.

To better compensate for the existing heterogeneity, we propose domain migration loss converters that jointly guide the training of the trunk network through the output of the framework's auxiliary branches and the labeling information of different domains. To achieve real-time extraction of finger vein patterns, we design two lighter and more efficient lightweight segmentation networks as the trunk and auxiliary branches in the Lite-HDNet framework for the purpose of more precise and rapid segmentation. Our major contributions are summarized below:

1) In this research, for the purpose of extracting finger vein features, we develop a brand-new, lightweight, domain-adaptive real-time segmentation system called Lite-HDNet. The framework can combine multiple datasets for training and use knowledge refinement to migrate feature knowledge from multiple finger vein datasets into the backbone network to improve the generalizability of the backbone network and weaken the heterogeneity existing among data. To the best of our knowledge, this is the first domain adaptation work in the field of finger vein extraction.

2) The middle-layer guided multi-domain feature knowledge transfer (MDFKT) technique is a new learning paradigm that we suggest. This method approximates steering the trunk network by using the feature distribution of finger veins from various domains. The trunk network can then concentrate on learning generic representations of data from several domains, enhancing model performance and generalizability.

3) The domain migration loss converter (DMLC) is proposed to better reduce the heterogeneity between different finger vein datasets and improve the generalizability of the model. This strategy combines the output of the auxiliary branch with the segmentation loss of the corresponding real labels, while the logits of the auxiliary branch and the trunk branch are jointly used to guide the training of the trunk network.

4) To achieve real-time extraction of finger vein patterns, we designed two lightweight networks for building the trunk branch and auxiliary branch in Lite-HDNet. On four-finger vein public datasets, the designed trunk and auxiliary networks achieve SOTA performance when compared to other lightweight segmentation networks.

The strategy proposed in this study produces the most cutting-edge experimental results on four publicly accessible finger vein datasets (SDU [6], MMCBNU [7], HKPU [8], and UTFVP [9]). On the NVIDIA integrated terminals,

we additionally test our method's real-time performance (JETSON NANO, JETSON TX2, JETSON XAVIER NX, JETSON AGX XAVIER).

The structure of this paper is built as follows: Section II presents related work in the field, while section III describes our proposed methodology. Then, section IV lists our experimental results, and Section V concludes the article.

II. RELATED WORK

A. METHOD FOR EXTRACTING FINGER VEINS BASED ON DEEP LEARNING

Finger vein feature extraction techniques based on deep learning have produced some outstanding outcomes. Qin et al. [10] proposed that convolutional neural networks (CNN) should be used to extract vein patterns and fully convolutional networks (FCN) should be applied to resume the missing finger vein patterns in segmented images. Fang et al. [11] proposed a dual-stream lightweight network to integrate raw finger vein images and mini ROI for efficient finger vein verification. Yang et al. [12] used a generative adversarial network (GAN) approach for finger vein pattern extraction and designed an adversarial training strategy and loss function to enhance the segmentation effect. In order to extract vein features, Qin et al. [13] suggested a unique network that combines the long-short-term memory model (LSTM) and CNN, where CNN learns robust features of vein texture pattern representation and LSTM retains complicated spatial relationships of vein patterns. With the use of a state-of-the-art semantic segmentation network, Jalilian et al. [14] extracted vein patterns from finger images and regarded them as actual vein features in biometric finger vein recognition. Zeng et al. [15] suggested an end-to-end approach for extracting vein textures that combines a fully convolutional neural network (FCN) with a conditional random field (CRF) that can capture complicated vein structural elements by adaptively modifying the receptive field based on the vein's scale and form. In the work of Noh et al. [16], they devoted themselves to extracting actual finger vein patterns directly from finger images without any pre- or post-processing assisted by a semantic segmentation convolutional neural network. Li et al. [17] established a finger vein infrared image segmentation dataset and proposed to segment finger vein images using the pyramidal structure and attention mechanism. However, the issues of weak model generalizability and subpar real-time performance have not been addressed in any of the aforementioned deep learning-based finger vein extraction projects. In this research, we offer a way to overcome these two drawbacks.

B. SOLVE THE PROBLEM OF DATA HETEROGENEITY BETWEEN DIFFERENT DOMAINS

Researchers in the field of medical image processing have compensated for the scarcity of medical images by federating multiple datasets, but the inherent heterogeneity of different datasets adds difficulties to the combined training among them. Many recent studies have used domain adaptation methods to deal with the heterogeneity that exists among domain data. To optimize the automatic grading function of breast cancer in lymph nodes, Wollmann et al. [18] adopted cyclic consistent generative adversarial networks (CycleGAN) as well as densely connected deep neural networks for domain adaptation methods. Javanmardi et al. [19] trained CNN with adversarial-based training methods to reduce the discrepancies between source domain and target domain datasets. Chen et al. [20] worked out a collaborative image and feature adaptation approach to effectively address the domain transfer problem by implementing adaptive co-fusion from the perspective of images and features. Wang et al. [21] established a low-rank representation-based multisite adaptation framework to find a common low-rank representation for data from multiple sites and reduce data heterogeneity between the target and source domains. Liu et al. [22] used multisite guided knowledge migration to improve the kernel and develop a domain-specific batch normalization layer to allow the network to estimate statistics and perform feature normalization for each site separately. Zhao et al. [23] refine the source classifier by minimizing the empirical Wasserstein distance between the source and the target, mapping the target to each source's feature space separately, and selecting source training samples closer to the target. Zhou et al. [24] made use of nonlinear transformations to enhance source-similar and source-different images and suggested the batch normalization layer in the proposed dual normalization-based model be normalized separately. He et al. [25] proposed a bidirectional image synthesis and segmentation module that used an intermediate data distribution generated for both domains to segment brain tumors. Han et al. [26] created a depth-symmetric architecture for medical image segmentation by introducing bidirectional alignment via a shared encoder and two private decoder schemes to simultaneously align the features between the source and target domains to narrow domain differences. These domain adaptation methods can combine multiple datasets to train the network, which improves model performance and generalizability; however, no domain adaptation method exists for finger vein extraction, and the method proposed in this paper is the first lightweight domain adaptation segmentation framework for finger vein extraction.

C. LIGHTWEIGHT NEURAL NETWORKS

Due to the unrealistic requirements of long training times and powerful computation hardware devices in deep neural networks, designing lightweight neural network models becomes an ideal solution to this problem. Hence, the design of lightweight neural network models has received considerable attention in academia and industry, with several lightweight models proposed. To separate the correlation in two directions, Xception [27] departs from spatial convolution and channel convolution. SqueezeNet [28] divides the convolution layer into an extension layer and a compression layer, with the compression layer compressing the number of

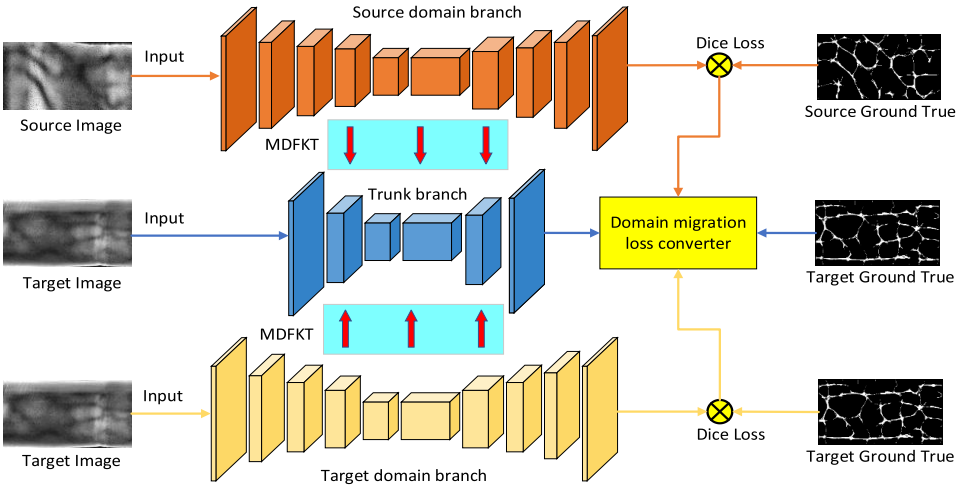

FIGURE 1. Our proposed lightweight domain-adaptive real-time segmentation framework (Lite-HDNet).

channels in the model. MobileNet [29] employs deep separable convolution to reduce model complexity while improving performance, and MobileNetV2 [30] proposes a reverse residual block to replace positive residuals. ShuffleNet [31] introduced the channel shuffle method to solve the problem of exchanging information between groups in grouped convolution. ShuffleNetV2 [32] discussed four basic principles for faster network operation by considering a real-time model. GhostNet [33] proposed a ghost module to compress the model by considering feature redundancy. MicroNet [34] proposed microfactor decomposition convolution to reduce node connectivity, resulting in a lower FLOP model, and introduced dynamic activation functions to improve nonlinearity.

III. PROPOSED METHOD

In this section, we give an overview of our proposed lightweight domain-adaptive real-time segmentation framework (Lite-HDNet) for finger vein pattern extraction, which is illustrated in Fig. 1. It is made up of a trunk branch and two auxiliary branches: the source domain branch and the target domain branch. The target domain finger vein images are used to train the trunk branches and target domain branches, while the source domain finger vein images are used to train the source domain branches. And the feature knowledge distribution extracted in the network layer of auxiliary and trunk branches is approximated by feature loss, which enhances the segmentation performance of the shallow trunk branches as well as their generalizability.

A. EQUATIONS

We propose an intermediate layer-guided multi-domain feature knowledge transfer strategy (MDFKT) to overcome the heterogeneity in finger vein data and improve the robustness of the trunk branch. The MDFKT strategy was influenced by research like that found in [35], [36], [37], [38], and [39], which introduced intermediate representations to create more accurate models of students and directly linked teacher and student knowledge of characteristics. The MDFKT strategy differs from all previous feature knowledge-based distillation strategies in that the trunk branch must learn not only the deep feature representation of the target domain finger vein data from the target domain branch but also the generic feature representation of the source domain finger vein data from the source domain branch, as well as how the trunk branch should trade off feature knowledge between domain branches, which we view as a challenge. Previous approaches to multi-teacher distillation have involved transferring knowledge by identifying models of teachers who perform well across multiple teachers or using a general averaging strategy that takes all teachers' knowledge into account equally. However, individual strengths may exist among teachers, and focusing solely on the knowledge of the best teacher may deplete the student model of the strengths of other teachers, whereas learning knowledge from all teachers equally may lead to conflicts among teachers during the knowledge transfer process, failing to optimize the best student model. Compared with the previous distillation work, which trained both teacher and student models with the same dataset, we proposed the MDFKT strategy, the first multi-domain multi-teacher distillation strategy, and this work is the first attempt at finger vein segmentation.

As previously stated, when solving the multi-domain multi-teacher distillation problem, we try to improve the trunk branch's generalization ability to the source domain data and its robustness in the hope of the trunk branch being able to accept the auxiliary branch's guidance to the greatest extent. Figure 2 depicts the MDFKT strategy's implementation. First, we use converters to resize the middle layer feature map of the source branch and the target branch so that they coincide with the middle layer feature map of the

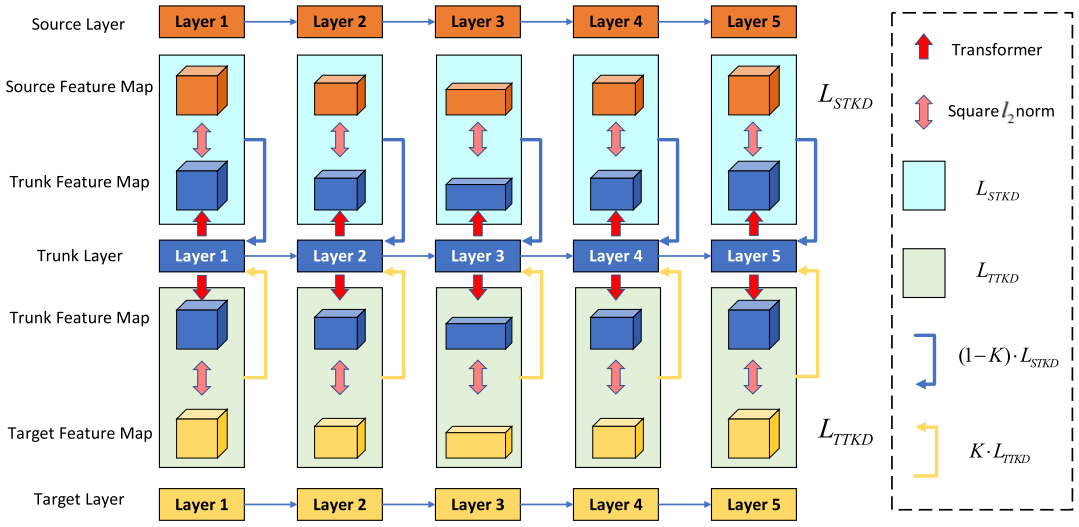

FIGURE 2. An overview of the multi-source feature knowledge transfer strategy (MDFKT) guided by the middle layer. The trunk branch's feature maps are aligned with the two auxiliary branches' feature map dimensions, and the corresponding feature loss functions aid the trunk branch in learning feature knowledge from different domains in a focused manner.

trunk branch, and then we measure the similarity between the two groups of feature maps with the help of the squared parametric as a distance measure. Finally, branch controllable variables K are introduced to control the share of the source domain branch's characteristic distillation loss and the target domain branch's characteristic distillation loss in the trunk branch guidance in order to obtain more robust and generalized trunk branches. The MDFKT strategy can be expressed by the following equation:

$$L_{STKD} = \sum_{i=1}^{n} Dist_{l_2}(Trans(F_M^i), F_s^i), i \in [1, n] \quad (1)$$

$$L_{TTKD} = \sum_{i=1}^{n} Dist_{l_2}(Trans(F_M^i), F_T^i), i \in [1, n] \quad (2)$$

where $Dist_{l_2}(,)$ denotes the similarity of two sets of feature maps measured using the square l_2 parametric; $Trans(\cdot)$ represents the converter to adjust the size of the feature map; F_M^i stands for the feature map of the trunk branch's middle layer; i refers to the middle layer of the i layer; and n shows the number of layers of the middle layer, which in this paper is $n = 5$. By substituting the square paradigm, Eq. (1) and Eq. (2) can be simplified to Eq. (3) and Eq. (4), respectively.

$$L_{STKD} = \frac{1}{2} \sum_{i=1}^{n} \left\| F_s^i - Trans(F_M^i) \right\|^2 \quad (3)$$

$$L_{TTKD} = \frac{1}{2} \sum_{i=1}^{n} \left\| F_T^i - Trans(F_M^i) \right\|^2 \quad (4)$$

At length, with the introduction of controllable variables, the overall loss of MDFKT can be expressed as:

$$L_{MSFKT} = (1 - K) \cdot L_{STKD} + K \cdot L_{TTKD}, K \in (0, 1) \quad (5)$$

The MDFKT strategy trains the trunk branch commonly by two auxiliary branches from different domains, so that the trunk branch can learn specific knowledge of different domains more comprehensively for the improvement in robustness and generalization ability of the trunk branch.

B. DOMAIN MIGRATION LOSS CONVERTER (DMLC)

In this paper, it is one of our objectives to use finger vein data from different domains to provide a more efficient training for the trunk network, to address the heterogeneity that exists between data from different finger veins and improve the generalization ability of the model. Therefore, it is inadequate to solely rely on the MDFKT strategy to guide trunk branch training. Previous works on image segmentation have shown that true labels are indispensable in guiding the training of segmentation networks, which inspires us to explore the effectiveness of employing true labels of finger veins from various domains to guide the training of trunk branches. Nevertheless, the real labels of different domains cannot directly guide the training of the trunk branch, which could lead to errors in the training results. we design a domain migration loss converter is introduced to indirectly guide the training of the trunk branch using the real labels of finger veins from different domains to further solve the heterogeneity in the data and improve the trunk branch's performance.

In essence, the domain migration loss converter is a training loss function that converts the trunk branch. We use the output of the auxiliary branch to match the corresponding true labels, then pass the label information of the source and target domains through the loss function, and finally combine the loss from the auxiliary branch with the trunk branch's segmentation loss to obtain the final loss for guiding trunk branch training. The design of the domain migration loss converter can be found in Fig. 3. It begins with denoting the output logits of the auxiliary branches and the trunk branch as a^s, a^t and a^m, respectively, with the minimization of discrepancy between the auxiliary branch output and the trunk branch output by the vanilla KD loss. This process can be expressed as Eq. (6) and Eq. (7).

$$L_{S-vanilla} = H\left(p^m, p^s\right) = H\left(\sigma\left(a^m; T\right), \sigma\left(a^s; T\right)\right)$$

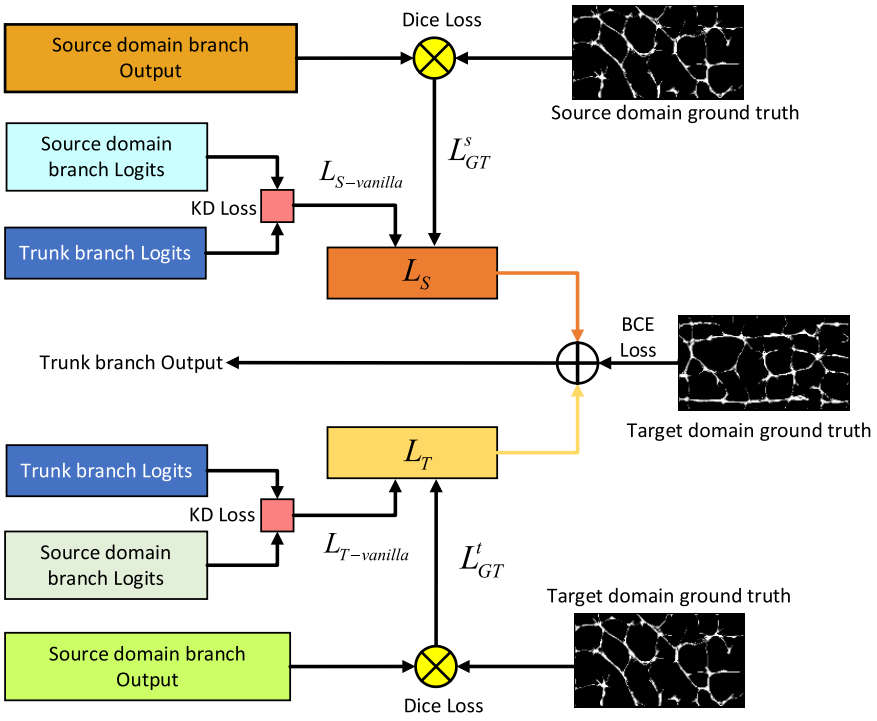

FIGURE 3. Domain migration loss converter.

$$= -\sum_{k=1}^{K} p^s[k]\, log p^m[k]$$
$$= - < p^s, log\ p^m > \quad (6)$$

$$L_{T-vanilla} = H\left(p^m, p^t\right) = H\left(\sigma\left(a^m; T\right), \sigma\left(a^t; T\right)\right)$$
$$= -\sum_{k=1}^{K} p^t[k] log p^m[k]$$
$$= - < p^t, log\ p^m > \quad (7)$$

where p^m is the trunk branch's output vector, p^s and p^t are the source and target branches' output vectors, respectively, T is the temperature of the softened logits, $\sigma(\cdot)$ is the softmax operation, $p[k]$ is the k component of the vector p, and $<,>$ is the vector's inner product operation.

The next step is to measure the similarity between the output of the auxiliary branch and the corresponding real labels, whose results will be utilized to adjust the Dice loss so as to obtain the loss function guided by the real labels from different domains to facilitate the main ranch training.

$$L^s_{GT} = \frac{\sum_i^N Dist(M_s, M_{GT}) \cdot 2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N Dist(M_s, M_{GT}) \cdot \sum_i^N g_i^2} \quad (8)$$

$$L^t_{GT} = \frac{\sum_i^N Dist(M_T, M_{GT}) \cdot 2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N Dist(M_T, M_{GT}) \cdot \sum_i^N g_i^2} \quad (9)$$

where p_i stands for the predicted binary segmentation pixel and g_i represents the label binary pixel, for a total of N pixels.

The loss function L_{GT} that represents the true labels of different domains and the vanilla KD loss function that reduces the discrepancy between the auxiliary branch and the trunk branch are then input to the loss converter in Fig.3. After weighted and averaged, they are used to guide the network training in the trunk branch. The loss function's operation in the loss converter can be written as follows

$$L_S = \frac{1}{2}(L_{S-vanilla} + L^s_{GT}) \quad (10)$$

$$L_T = \frac{1}{2}(L_{T-vanilla} + L^t_{GT}) \quad (11)$$

At last, we use the loss functions L_S and L_T output by the loss converter with the binary BCE loss for guiding the training of the trunk branch. To control the dependence of the trunk branch on the knowledge of different domains, the hyperparameter γ was introduced to regulate the loss guidance of the source domain branch and the target domain branch, and the final loss function used to guide the trunk branch training can be expressed as

$$L_{DM} = \frac{1}{3}((1-\gamma)\cdot L_S + \gamma \cdot L_T + L_{BCE}) \quad (12)$$

In short, our proposal of the domain migration loss converter serves to enable the trunk branch to better learn specific knowledge from different domains and resolve heterogeneity between different finger vein data, which will be more useful in improving the model's generalization ability.

C. DESIGN OF LIGHTWEIGHT SEGMENTATION MODEL

To achieve more efficient finger vein pattern extraction and improve the real-time performance of the model, we designed two lightweight segmentation networks to build Lite-HDNet, which are the trunk branch network and the auxiliary branch

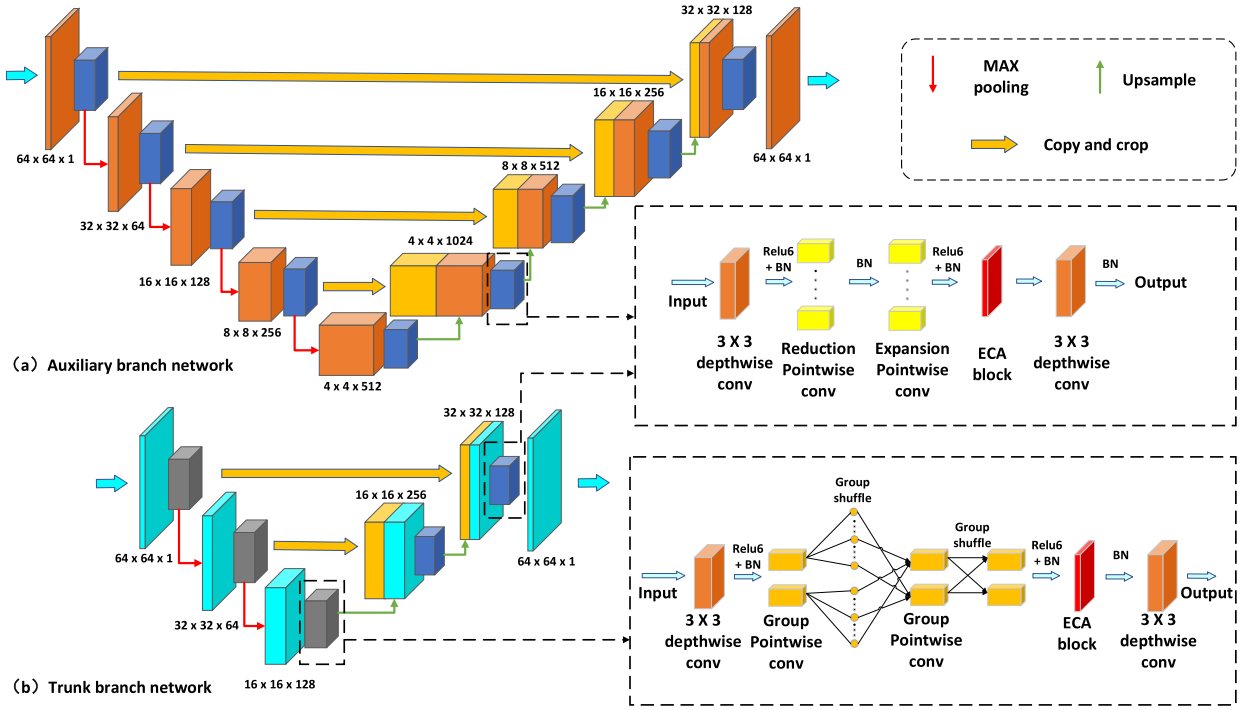

FIGURE 4. The structures of the auxiliary branch network and the trunk branch network.

TABLE 1. Details of finger vein dataset.

Datasets	Finger number	Total number of images	Size of original images
SDU [5]	636	3816	320×240
MMCBNU [6]	600	6000	640×480
HKPU [7]	312	3744	513×256
UTFVP [8]	360	1440	672×380

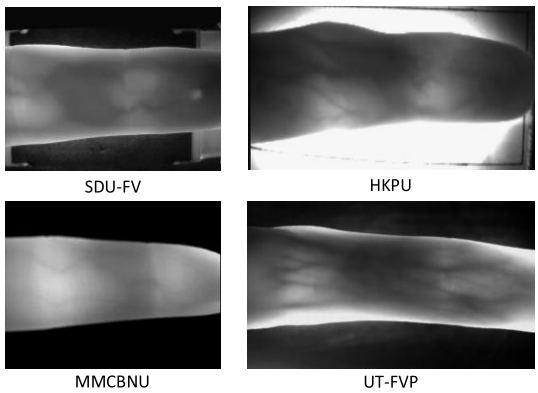

FIGURE 5. Raw images of the four finger vein datasets.

network, and the network structures are shown in Fig. 4. We use the Unet [40] model of architecture as the benchmark for the design, which is because its U-shaped structure can take into account both global and local features of the image, while the shallow part of the network focuses more on local features such as texture, and the deep part of the network focuses more on the essential features of the image. We use lightweight depth-separable convolution to replace the standard convolution and follow the principle of spatial transformation in higher dimensions, which effectively alleviates the loss of feature information and gradient confusion.

To enhance the effectiveness of the module, we added the attention module ECA [41] to avoid performing dimensionality reduction operations in the attention dimension, which in turn affects the network performance.

To make the trunk branch network have higher real-time performance in the final prediction, we use the idea of grouping in the trunk network to reduce the connectivity between point convolution channels and group the sparse point convolution. By sparsifying the connectivity between point convolutions, we can reduce the complexity of convolutional operations and introduce two different group shuffle strategies for inter-group information exchange to ensure the performance of the model.

Compared with previous lightweight segmentation networks, our designed trunk branch network and auxiliary branch network possess better segmentation performance and fewer parameters, with experimental results shown in Table 3.

IV. EXPERIMENTS

A. DATASETS AND EXPERIMENTAL CONFIGURATION

The finger vein dataset SDUMLA-HMT (SDU) from Shandong University's Machine Learning and Data Mining Laboratory, the finger vein dataset MMCBNU 6000

(MMCBNU) from Jeonbuk National University, Korea, the finger vein dataset from Hong Kong Polytechnic University (HKPU), and the finger vein dataset from the University of Twente, the Neth, were chosen separately to validate the effectiveness of the proposed method. The imaging effect varies greatly due to the different acquisition devices they used, and the placement of the fingers, the area of the collected fingers, and the clarity of the blood vessels in the fingers are all different, as shown in Fig. 5. To be fair, we performed the same preprocessing operation in the image preprocessing stage for all four datasets, and since the inconsistent image sizes in the datasets may affect the fairness of the experimental results, we resized the images and labels of the four finger vein datasets after preprocessing to a uniform size of 270 × 150. In this study, we used the maximum curvature method [3] to generate true labels for the four datasets of finger veins for experiments and divided the data in each dataset into training and test sets in a 4:1 ratio.

Our experiments were run on the pytorch1.8.0 framework, which employed a single NVIDIA GeForce RTX 3090 TURBO GPU for training, a SGD optimizer for network training, a momentum of 0.9, a weight decay of 0.01, and a batch size of 256. During training, the chunking strategy was adopted for feature extraction and prediction, which not only completed the work of expanding data but also better predicted the finger vein detail part. We set the step size of each patch's width and height to 5, extract multiple consecutive overlapping blocks for each image, and calculate the probability that a pixel is a vein by averaging the probability of all predicted blocks that cover the pixel. Moreover, we validate the real-time performance of finger vein pattern extraction on NVIDIA embedded terminals JETSON NANO, JETSON TX2, JETSON XAVIER NX, and JETSON AGX XAVIER.

B. EVALUATION INDICATORS

In this paper, we used three segmentation metrics—Dice, AUC, and accuracy—as the basis for evaluating the performance of finger vein segmentation. Dice: The most frequently used segmentation metric, which represents the ratio of the intersection area of two segmentation results and labeling to the total area.

C. ABLATION EXPERIMENTS

In this section, we conduct a large number of experiments to demonstrate the effectiveness of our proposed method. In subsection (a), the effectiveness of the MDFKT and domain migration loss converter (DMLC) proposed in this paper is verified, i.e., finger vein images from other domains can be used to improve the finger vein pattern extraction from the target domain. In subsection (b), the segmentation results of the two designed lightweight segmentation networks and some recent lightweight segmentation networks on the finger vein dataset are analyzed to demonstrate their feasibility. In subsection (c), the real-time performance of the proposed two lightweight segmentation networks is tested on the PC side and on NVIDIA embedded terminals. Finally, we verify in subsection (d) that the Lite-HDnet framework proposed in this paper is a lightweight and efficient framework for finger vein segmentation.

(a)Effectiveness of MDFKT and Domain Migration Loss Converters (DMLC): To validate the effectiveness of MDFKT and DMLC, we verified the effectiveness of the proposed methods on four finger vein datasets, respectively, and the experimental results are shown in Table 2. Among them, we used the Unet network as the baseline and used the control variables method to validate the effectiveness of each of the MDFKT and DMLC methods as well as the combined effect of the two methods. In addition, since one of the main objectives of this paper is to improve finger vein pattern extraction in the target domain using finger vein images from different domains, we introduce finger vein images from different domains as source domains in the experiments in Table 2 to assist the training.

We have selected different source-domain finger vein images for their auxiliary training and observed the experimental results to find that the MDFKT and DMLC methods can combine the source-domain finger vein images to improve the segmentation effect of Unet. When the target domain is SDU, the MDFKT method improves the model performance and has a greater effect on the improvement of the AUC metric and a smaller improvement of the Dice metric, while the DMLC method can better improve the Dice value, and the improvement of the AUC metric is slightly smaller than the increase of the MDFKT method. When trained with both MDFKT and DMLC, the overall experimental results are better than those of the single method, and the segmentation effect is substantially improved compared with the baseline. By analyzing the results of different source domains, it can be found that introducing different source domains for assisted training results in improved segmentation of the target domain, even for the basic Unet. When combined with the MDFKT and DMLC methods, the segmentation performance of the network can be improved to the greatest extent. Our method motivates the network to better learn the finger vein features from different source domains and make full use of the source domain features to improve the extraction effect of the target domain finger vein pattern. In addition, by comparing different source domain images, it can be found that the segmentation effect is improved the most when UTFVP is selected as the source domain. Thus, it can be concluded that the data heterogeneity between UTFVP and SDU is the least compared to other finger vein datasets, which is more beneficial to improving the segmentation effect of SDU.

Similar to the conclusion obtained when the target domain is SDU, when the target domain is MMCBNU, the segmentation effect of the target domain can be improved by using MDFKT, DMLC, and the joint method, among which the best results are obtained by the joint method. By comparing the results of different source domains, we found that the segmentation effect of the base Unet is optimal when the source domain is MMCBNU, while the direct introduction of other source domains leads to a decrease in the segmentation

TABLE 2. Comparison of segmentation performance of MDFKT and DMLC methods with baseline Unet.

Target domain	Method	Source domain							
		SDU		MMCBNU		HKPU		UTFVP	
		Dice	AUC	Dice	AUC	Dice	AUC	Dice	AUC
SDU	Unet	46.82%	85.77%	48.98%	86.01%	49.59%	88.19%	50.27%	88.62%
	Unet+MDFKT	46.88%	87.63%	49.06%	88.53%	49.60%	89.48%	50.38%	90.32%
	Unet+DMLC	48.14%	87.24%	49.21%	88.26%	49.88%	89.11%	51.13%	89.87%
	Unet+M+D	48.43%	**88.35%**	**49.39%**	**89.14%**	**49.95%**	**90.11%**	**51.31%**	**91.02%**
MMCBNU	Unet	48.45%	88.67%	52.24%	92.16%	50.23%	90.08%	50.34%	90.12%
	Unet+MDFKT	50.23%	92.33%	52.36%	93.42%	51.17%	93.77%	51.03%	93.63%
	Unet+DMLC	52.57%	91.71%	53.11%	92.58%	52.88%	92.78%	53.25%	92.89%
	Unet+M+D	**53.49%**	**94.06%**	**53.32%**	**93.80%**	**53.68%**	**94.54%**	**53.73%**	**94.17%**
HKPU	Unet	52.28%	89.00%	53.61%	90.15%	56.49%	91.73%	55.83%	91.01%
	Unet+MDFKT	54.97%	92.45%	54.76%	93.12%	56.44%	93.63%	56.23%	93.76%
	Unet+DMLC	56.46%	92.01%	56.83%	92.27%	**56.60%**	92.11%	57.10%	92.98%
	Unet+M+D	**57.16%**	**93.74%**	**57.39%**	**94.02%**	56.52%	**93.72%**	**57.45%**	**94.22%**
UTFVP	Unet	50.92%	86.32%	51.54%	88.34%	51.91%	89.16%	56.54%	90.57%
	Unet+MDFKT	52.13%	91.35%	52.66%	91.49%	53.28%	91.80%	56.49%	91.84%
	Unet+DMLC	55.49%	90.71%	54.88%	90.95%	55.68%	91.31%	56.60%	91.00%
	Unet+M+D	**56.81%**	**92.41%**	**56.95%**	**92.49%**	**57.29%**	**92.68%**	**56.66%**	**92.13%**

effect of MMCBNU. This is due to the fact that the baseline method does not use any domain adaptation method to address the inherent heterogeneity between different finger vein images, making it difficult for the model to learn a better representation from multiple datasets, and the results in Table 2 demonstrate that it is difficult to obtain a more robust segmentation effect by directly uniting multiple finger vein datasets for training. The MDFKT and DMLC methods suggested in this paper can effectively address the issue. The heterogeneity between data sets can be gradually reduced by using one method alone, and the best results can be obtained by combining the two methods.

The analysis shows that when the source domain is SDU, Dice and AUC improve by 5.04% and 5.39% when both methods are used simultaneously compared with the baseline; when the source domain is MMCBNU, Dice and AUC improve by 1.08% and 1.64%; when the source domain is HKPU, Dice and AUC improve by 3.45% and 4.46%; and when the source domain is UTFVP, Dice and AUC improve by 3.39% and 4.05%. The effectiveness of the proposed method is further verified by analyzing the experiments with the target domain as MMCBNU.

When the target domain is HKPU and UTFVP, the same conclusions are obtained as when the target domain is MMCBNU. The baseline Unet results in a degradation of segmentation performance on the target domain when the target domain is directly trained jointly with the source domain images, and the degradation of performance varies depending on the degree of heterogeneity differences between the datasets. In the experiments where the target domain is HKPU, we can find that the source domain selection of SDU has the greatest effect on the segmentation effect of the target domain, while the source domain selection of UTFVP has the least effect on the segmentation effect, which can infer that the heterogeneity between UTFVP and HKPU is smaller than that between SDU and HKPU. When the target domain is UTFVP, the heterogeneity between HKPU, MMCBNU, and UTFVP is much smaller than that of SDU. The proposed method can not only solve the problem that heterogeneity between source and target domains leads to the degradation of finger vein segmentation but also improve finger vein pattern extraction in the target domain by learning the feature representation of the source domain so that the segmentation performance can be improved.

In this subsection, the experimental results show that our proposed method can effectively use different finger vein datasets to improve the model segmentation performance and enhance the extraction of target domains, and it can well solve the heterogeneity problem between different finger vein datasets to enhance the generalizability of the model.

(b)The designed lightweight segmentation network's robustness: In order to improve the segmentation of finger vein patterns and ensure a lightweight model with high real-time performance, we designed two effective lightweight segmentation networks: a trunk branching network and an auxiliary branching network. In this subsection, we experimentally compare the two proposed networks with some lightweight segmentation networks to verify the effectiveness of the proposed networks.

We compare the performance of the designed trunk branch network and auxiliary branch network with some lightweight segmentation networks on four publicly available finger vein datasets, and the results are shown in Table 3. According to the experimental results, our proposed auxiliary branch network outperforms Unet and other lightweight segmentation networks on all datasets. It is worth noting that the auxiliary

TABLE 3. Performance of our proposed Lightweigjt network on NIR finger vein datasets.

Dataset	Network	Dice	AUC	Params	Flops	Mult-Add
SDU	Unet	46.82%	85.77%	34.52M	4.09G	8.17G
	Squeeze Unet	48.74%	84.40%	2.893M	296.05M	586.44M
	Mobile Unet	48.32%	86.97%	3.932M	498.13M	985.04M
	Shuffle Unet	49.34%	88.05%	516K	68.27M	129.82M
	Ghost Unet	48.99%	89.34%	6.432M	130.11M	259.36M
	Auxiliary branch	**50.06%**	**90.14%**	516K	42.97M	83.73M
	Trunk branch	48.68%	89.97%	**60.59k**	**28.25M**	**54.49M**
MMCBNU	Unet	52.24%	92.16%	34.52M	4.09G	8.17G
	Squeeze Unet	49.22%	90.34%	2.893M	296.05M	586.44M
	Mobile Unet	52.47%	91.32%	3.932M	498.13M	985.04M
	Shuffle Unet	52.15%	91.40%	516K	68.27M	129.82M
	Ghost Unet	52.73%	93.54%	6.432M	130.11M	259.36M
	Auxiliary branch	**53.94%**	**94.15%**	516K	42.97M	83.73M
	Trunk branch	52.57%	93.62%	**60.59k**	**28.25M**	**54.49M**
HKPU	Unet	56.49%	91.73%	34.52M	4.09G	8.17G
	Squeeze Unet	56.98%	90.71%	2.893M	296.05M	586.44M
	Mobile Unet	56.45%	91.31%	3.932M	498.13M	985.04M
	Shuffle Unet	56.67%	92.26%	516K	68.27M	129.82M
	Ghost Unet	57.35%	93.15%	6.432M	130.11M	259.36M
	Auxiliary branch	**58.29%**	**94.88%**	516K	42.97M	83.73M
	Trunk branch	56.87%	93.88%	**60.59k**	**28.25M**	**54.49M**
UTFVP	Unet	56.54%	90.57%	34.52M	4.09G	8.17G
	Squeeze Unet	55.34%	90.19%	2.893M	296.05M	586.44M
	Mobile Unet	56.00%	90.35%	3.932M	498.13M	985.04M
	Shuffle Unet	56.08%	91.72%	516K	68.27M	129.82M
	Ghost Unet	56.57%	92.39%	6.432M	130.11M	259.36M
	Auxiliary branch	**57.96%**	**93.92%**	516K	42.97M	83.73M
	Trunk branch	55.84%	92.89%	**60.59k**	**28.25M**	**54.49M**

branch network has only 516K parameters and 42.97M flops, which is 25.3M fewer flops compared to Shuffle Unet. Compared with Unet, the auxiliary branch network not only has a more robust segmentation performance but also has 66 times fewer model parameters and 95 times fewer Flops. The proposed trunk branch network is smaller than the auxiliary branch network, with only 60.59K parameters and 28.25M loops. It is not only much smaller in terms of model parameters than other lightweight segmentation networks, but its segmentation performance is roughly comparable to and even exceeds that of some networks.

According to the results of Table 3, we designed two lightweight segmentation networks (a trunk branching network and an auxiliary branching network) that performed exceptionally well in segmenting finger vein patterns and modeling lightweighting. For visual comparison, we also present some examples of finger vein segmentation in Fig. 6. The extracted veins are smoother and more consistent with the vascular characteristics, and they also lessen the generation of burrs to some extent. The auxiliary branching network is more advantageous in extracting the fine details of the finger veins and illustrating the fine branches of the veins. The trunk branching network greatly ensures the model's segmentation performance while keeping the model lightweight, and the segmentation effect is superior to Unet and Mobile Unet.

In this subsection, we validate the segmentation results of the proposed two lightweight segmentation networks on finger vein images, and the results demonstrate the excellent performance of the trunk branching network and the auxiliary branching network in terms of both segmentation performance and model lightweighting.

(c)Verify the real-time nature of finger vein extraction: In this subsection, we investigate the real-time performance

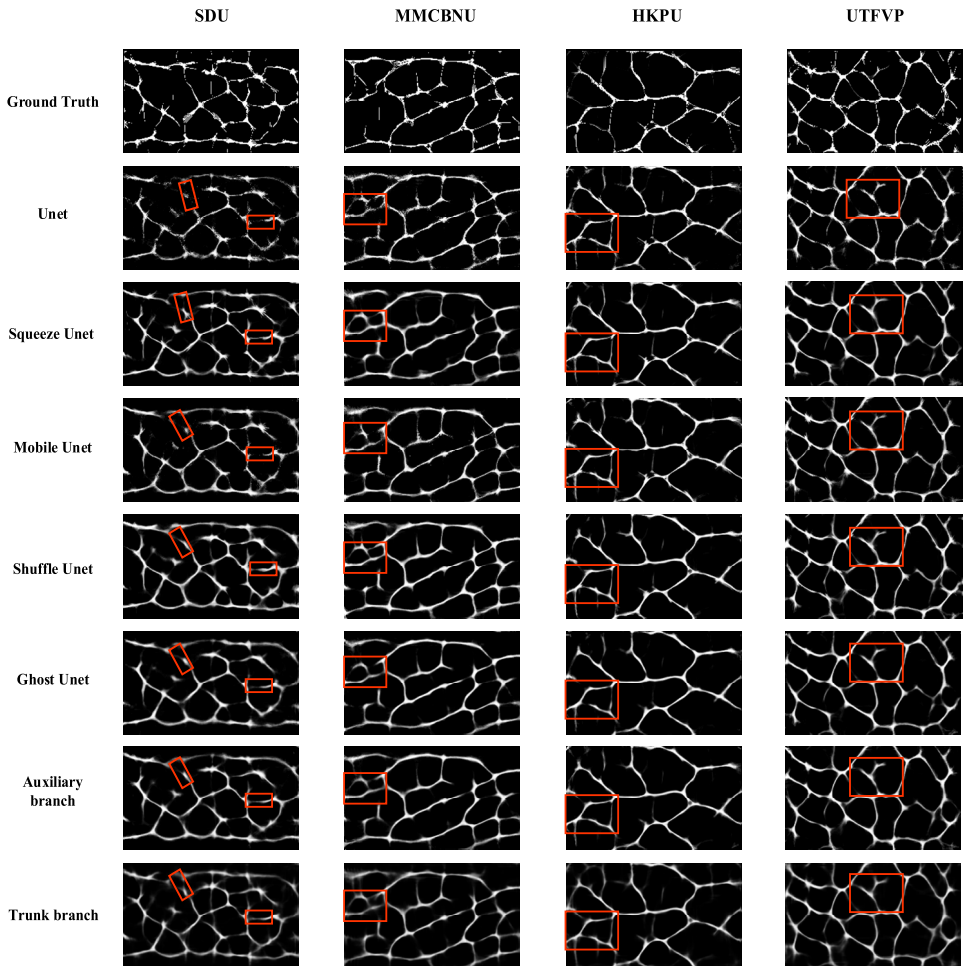

FIGURE 6. Results of finger vein extraction for the two lightweight networks we designed as well as other networks.

TABLE 4. Real-time experiment on the PC side.

Network	Time（s） / 1 picture				Testing Memory Usage
	SDU	MMCBNU	HKPU	UTFVP	
Squeeze Unet	0.115s	0.113s	0.107s	0.108s	2321MiB
Mobile Unet	0.080s	0.080s	0.079s	0.081s	2559MiB
Shuffle Unet	0.099s	0.095s	0.100s	0.104s	2325MiB
Ghost Unet	0.071s	0.072s	0.070s	0.071s	1915MiB
ours	**0.053s**	**0.054s**	**0.052s**	**0.054s**	**1859MiB**

of the proposed method on the PC side and the embedded side, respectively. Due to the low computing power of the embedded terminal, the number of images in a single test cannot be too large. To ensure fairness, we limit the number of images in each batch of tests to 20 on both PC and embedded terminals, and compare the average time required for inference on a single image by calculating the average time required for inference on a single image. Among them, due to the unsatisfactory computing power of Jetson NANO, which cannot support single processing of 20 finger vein images, we set the number of images per batch tested on Jetson NANO to 5 and obtained the time required for single image processing by summing and averaging. Table 4 shows the experimental results comparing the real-time performance of our proposed method with some lightweight networks on the PC side. The results indicate that our method outperforms general lightweight methods in terms of real-time performance, processing a single finger vein image in as little as 0.052s and requiring less memory when tested. When the experiments are conducted on the PC side, we note that the variation in the number of parameters and Flops of the model has less impact on the real-time performance of the network,

TABLE 5. Real-time experiment on the NVIDIA embedded terminals.

Embedded terminal	Network	Time（s） / 1 picture			
		SDU	MMCBNU	HKPU	UTFVP
JETSON NANO	Squeeze Unet	3.6250s	3.5699s	3.7136s	3.6293s
	Mobile Unet	3.4790s	3.5030s	3.5920s	3.5154s
	Shuffle Unet	2.7109s	2.6864s	2.8120s	2.7546s
	Ghost Unet	1.6275s	1.6000s	1.6447s	1.6229s
	ours	**0.8039s**	**0.7938s**	**0.7978s**	**0.8066s**
JETSON TX2	Squeeze Unet	0.6618s	0.6648s	0.6748s	0.6610s
	Mobile Unet	0.5243s	0.5176s	0.5430s	0.5163s
	Shuffle Unet	0.6778s	0.6611s	0.6805s	0.6661s
	Ghost Unet	0.4133s	0.4108s	0.4268s	0.4139s
	ours	**0.3291s**	**0.3277s**	**0.3362s**	**0.3303s**
JETSON XAVIAR NX	Squeeze Unet	0.4836s	0.4823s	0.5322s	0.4822s
	Mobile Unet	0.3977s	0.3988s	0.4213s	0.3824s
	Shuffle Unet	0.4210s	0.4077s	0.4393s	0.4149s
	Ghost Unet	0.3437s	0.3477s	0.3528s	0.3353s
	ours	**0.2420s**	**0.2417s**	**0.2624s**	**0.2542s**
JETSON AGX XAVIAR	Squeeze Unet	0.4060s	0.4081s	0.4188s	0.4078s
	Mobile Unet	0.3192s	0.3176s	0.3391s	0.3167s
	Shuffle Unet	0.3800s	0.3793s	0.3935s	0.3795s
	Ghost Unet	0.3022s	0.3029s	0.3366s	0.2951s
	ours	**0.2293s**	**0.2272s**	**0.2305s**	**0.2264s**

TABLE 6. Performance comparison of Lite-HDnet framework with baseline on finger vein datasets.

Datasets		Dice			AUC		
Source Domain	Target Domain	Unet	Unet+MDFKT +DLMC	Lite-HDnet	Unet	Unet+MDFKT +DLMC	Lite-HDnet
SDU	SDU	46.82%	48.43%	**50.12%**	85.77%	88.35%	**91.64%**
MMCBNU		48.98%	49.39%	**49.82%**	86.01%	89.14%	**91.78%**
HKPU		49.59%	49.95%	**50.38%**	88.19%	90.11%	**92.20%**
UTFVP		50.27%	51.31%	**52.15%**	88.60%	90.87%	**92.27%**
FUSION		51.74%	51.89%	**52.27%**	88.62%	91.02%	**92.52%**
MMCBNU	MMCBNU	52.24%	53.32%	**53.70%**	92.16%	93.80%	**94.22%**
SDU		48.45%	53.49%	**54.02%**	88.67%	94.06%	**94.41%**
HKPU		50.23%	53.68%	**54.34%**	90.08%	94.54%	**94.69%**
UTFVP		50.34%	53.73%	**54.35%**	90.12%	94.17%	**94.73%**
FUSION		51.87%	54.05%	**54.52%**	91.50%	94.68%	**94.88%**
HKPU	HKPU	56.49%	56.52%	**57.12%**	91.73%	93.72%	**94.68%**
SDU		52.28%	57.16%	**57.22%**	89.00%	93.74%	**94.71%**
MMCBNU		53.61%	57.39%	**57.49%**	90.15%	94.02%	**94.84%**
UTFVP		55.83%	57.45%	**57.54%**	91.01%	94.22%	**94.88%**
FUSION		56.52%	57.83%	**58.01%**	91.48%	94.58%	**95.01%**
UTFVP	UTFVP	56.54%	56.66%	**57.10%**	90.57%	92.13%	**93.32%**
SDU		50.92%	56.81%	**57.24%**	86.32%	92.41%	**93.45%**
MMCBNU		51.54%	56.95%	**57.58%**	88.34%	92.49%	**93.76%**
HKPU		51.91%	57.29%	**57.65%**	89.16%	92.68%	**93.83%**
FUSION		54.71%	57.60%	**58.07%**	89.75%	93.06%	**94.13%**

but the model's complexity and memory usage are the more important factors affecting the real-time performance of the model. Ma et al. [29] proposed that the key factor influencing model speed is the memory access cost (MAC) required for the model to run, whereas we discovered that model complexity can also have a significant impact on model speed. Table 5 shows the parameters, Flops, and multiplication and addition operations of the lightweight network, in which the model size and operations of Shuffle Unet are much smaller than those of Ghost Unet, but in the real-time experiments, the real-time performance of Shuffle Unet is worse than that of Ghost Unet, and the memory occupation is higher.

This is because there are too many group convolution and channel shuffling operations in Shuffle Unet, which complicate the network structure and reduce the real-time performance.

Table 5 lists the experimental results on the NVIDIA embedded terminals, from which our method achieves the best results on all four embedded terminals with different arithmetic powers, requires less inference time on the embedded terminal compared to other lightweight models, and achieves fast segmentation of finger vein images without using any quantization operations during the experiment. For JETSON NANO equipped with the lowest arithmetic power,

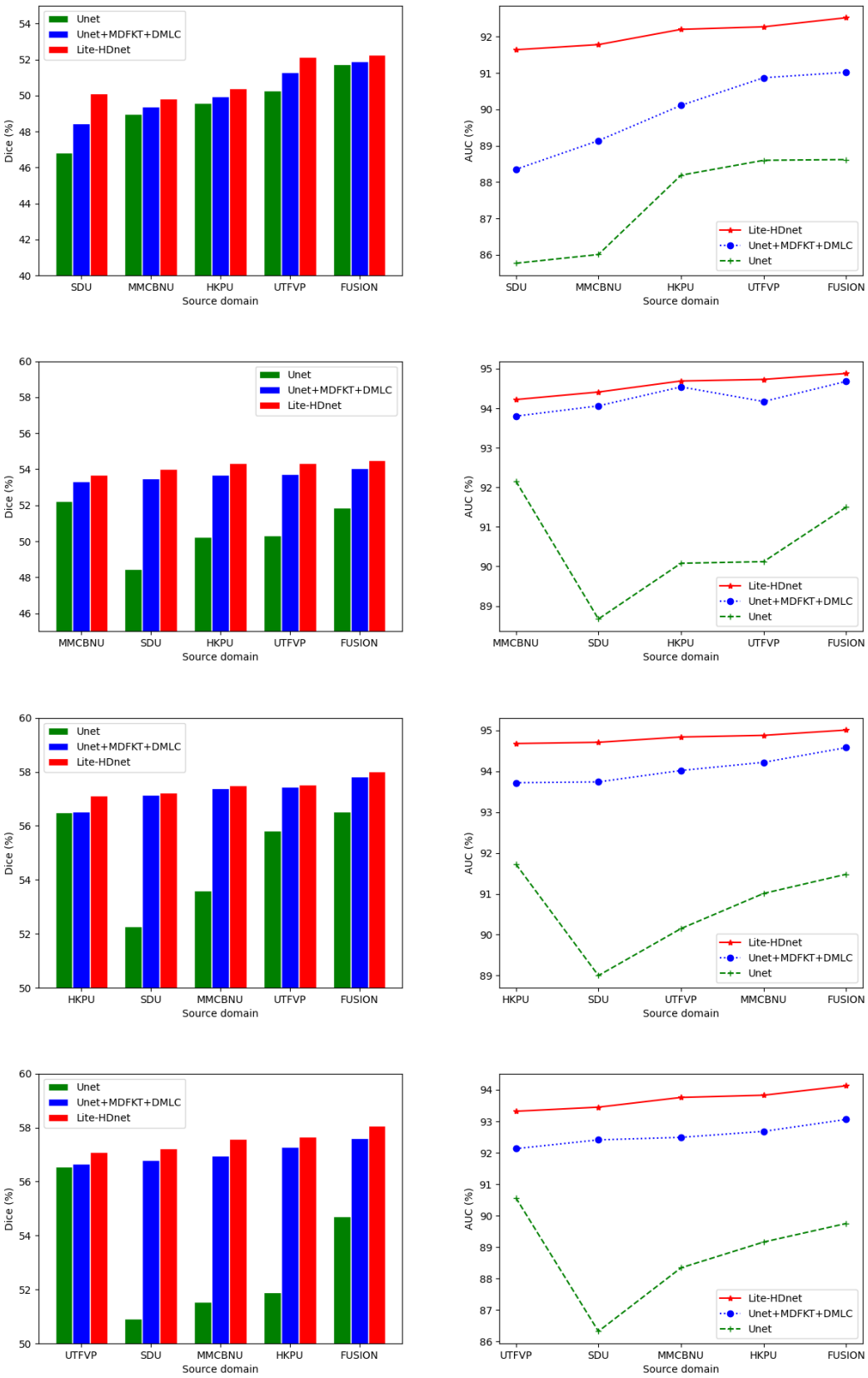

FIGURE 7. Visualization of the performance of the Lite-HDnet framework with Unet.

our proposed method only takes 0.7938s to extract a finger vein pattern image in the fastest situation.

(d)Analysis of the effectiveness of Lite-HDnet: In this subsection, we evaluate the Lite-HDnet that was created by combining the three techniques. The domain migration loss converter with the Lite-HDnet, which can be divided into four different groups of experiments depending on the target domain, is compared to the original Unet, the Unet using MDFKT and the domain migration loss converter. From the experimental results in Table 6, it can be seen that if the source domain finger vein images and the target domain finger vein images are trained jointly directly without the method proposed in this paper, there is heterogeneity between the datasets due to the different ways and devices used to collect different finger vein datasets, which leads to the degradation of the model training effect.

The proposed method can not only alleviate the heterogeneity between different datasets and improve the generalization performance of the model but also help the model learn more robust feature representation and richer feature information from other datasets so that it can extract finger veins better. When other finger vein datasets are introduced as the source domain, the model performance is better than when the same dataset is used as the source domain. The best model performance is achieved when a mixed dataset is used as the source domain.

When the target domain is SDU, direct introduction of the source domain finger vein image can improve the segmentation effect on the target domain, and our method not only fully learns the feature knowledge of different domains and transfers the feature knowledge to the target domain to improve the segmentation effect on the target domain, but also creates the model with high real-time performance to achieve real-time finger vein pattern segmentation. When the source domain is SDU, Lite-HDnet increases Dice by 3.3% and AUC by 5.87% compared to Unet; when the source domain is MMCBNU, Lite-HDnet increases Dice by 0.84% and AUC by 5.77% compared to Unet. When the source domain is HKPU, Lite-HDnet improves Dice by 0.79% and AUC by 4.01% compared to Unet; when the source domain is UTFVP, Lite-HDnet improves Dice by 1.88% and AUC by 3.9% compared to Unet; when the source domain is mixed dataset, Lite-HDnet improves Dice and AUC by 0.53% and 3.9% compared to Unet. Dice and AUC are improved by 0.53% and 3.9%, respectively. The difference between Lite-HDnet and Unet after using MDFKT and DMLC is that the segmentation network is different. Lite-HDnet is composed of the trunk branch network and the auxiliary branch network designed in this paper. From the experiments, we can observe that Lite-HDnet has better segmentation results compared with Unet using MDFKT and DMLC methods, thanks to the better performance of the two lightweight segmentation networks we designed and the fact that the Lite-HDnet model is lighter and has higher real-time performance.

When the target domain is MMCBNU, Unet directly introduces the source domain finger vein images with the target domain to jointly train the network, which does not improve the segmentation performance of the network. This is due to the different acquisition devices, imaging methods, and finger placement positions used in different datasets of finger vein images, resulting in a large heterogeneity between different finger vein datasets, and direct joint training will affect the extraction effect of finger veins and make the performance of the model drop sharply. Lite-HDnet is proposed to be a good solution to this problem. As shown in Table 6, when the source domain is SDU, Lite-HDnet improves Dice and AUC by 5.57% and 5.74% compared to Unet. When the source domain is MMCBNU, Lite-HDnet improves Dice by 1.46% and AUC by 2.06% compared to Unet. When the source domain is HKPU, Lite-HDnet improves Dice by 4.11% and AUC by 4.61% compared to Unet. When the source domain is UTFVP, Lite-HDnet improves Dice by 4.01% and AUC by 4.61% compared to Unet. When the source domain is FUSION, Lite-HDnet improves Dice by 2.65% and AUC by 3.38%. Compared with Unet after using MDFKT and DMLC methods, Lite-HDnet has better segmentation results, further proving that the two lightweight segmentation networks we designed are more effective for the finger vein segmentation task. Lite-HDnet improves Dice by 2.65% and AUC by 3.38%. Compared with Unet after using MDFKT and DMLC methods, Lite-HDnet has better segmentation results, further proving that the two lightweight segmentation networks we designed are more effective for the finger vein segmentation task.

When the target domain is HKPU and the target domain is UTFVP, the conclusions obtained from analyzing the experimental results in Table 6 are the same as those on MMCBNU, further proving that Lite-HDnet is effective and feasible. To make it easier to compare the performance of Lite-HDnet and the baseline, we plotted Fig. 7, in which the enhancement effect of Lite-HDnet can be observed more visually.

In this subsection, it is demonstrated through extensive experiments that our proposed Lite-HDnet can improve the extraction of finger vein patterns in the target domain using finger vein images from different domains. Lite-HDnet can effectively resolve data heterogeneity to ensure the model's segmentation effect.

V. CONCLUSION

In this paper, we propose a new lightweight domain adaptation real-time segmentation framework (Lite-HDNet) to solve the problems of poor generalizability and low real-time performance in current deep neural network-based finger vein pattern extraction. This is the first work to improve finger vein pattern extraction using a domain-adaptive approach to combine multiple finger vein datasets. This framework can learn common feature representations in multiple finger vein datasets with better segmentation performance and stronger generalizability. Our proposed multi-domain feature knowledge transfer strategy (MDFKT) and domain migration loss converter can effectively address the problem of poor generalizability of the trained models due to the heterogeneity

among data, and the effectiveness of the proposed method has been validated on four publicly available finger vein datasets (SDU, MMCBNU, HKPU, and UTFVP). We borrow two lightweight segmentation networks (the auxiliary branch network and the trunk branch network), giving Lite-HDNet better segmentation capability and higher real-time performance. The auxiliary branch network and trunk branch network are then tested against other lightweight segmentation networks, with the results indicating that our proposal has a lighter and more efficient segmentation network. Among them, the parameter of the auxiliary branch network is only 516K, and the parameter of the trunk branch network is only 60.59K. Furthermore, the real-time performance of the proposed method has been verified on the NVIDIA series of embedded terminals, which are capable of extracting finger veins in real time on the embedded terminals without quantization.

REFERENCES

[1] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Mach. Vis. Appl.*, vol. 15, no. 4, pp. 194–203, Oct. 2004.

[2] H. Qin, "Region growth–based feature extraction method for finger-vein recognition," *Opt. Eng.*, vol. 50, no. 5, May 2011, Art. no. 057208.

[3] N. Miura, A. Nagasaka, and T. Miyatake, "Extraction of finger-vein patterns using maximum curvature points in image profiles," *IEICE Trans. Inf. Syst.*, vols. 90, no. 8, pp. 1185–1194, Aug. 2007.

[4] J. Yang, J. Yang, and Y. Shi, "Finger-vein segmentation based on multi-channel even-symmetric Gabor filters," in *Proc. IEEE Int. Conf. Intell. Comput. Intell. Syst.*, vol. 4, Nov. 2009, pp. 500–503.

[5] B. Huang, Y. Dai, R. Li, D. Tang, and W. Li, "Finger-vein authentication based on wide line detector and pattern normalization," in *Proc. 20th Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 1269–1272.

[6] Y. Yin, L. Liu, and X. Sun, "SDUMLA-HMT: A multimodal biometric database," in *Biometric Recognition*, Beijing, China. Berlin, Germany: Springer, 2011, pp. 260–268.

[7] Y. Lu, S. J. Xie, S. Yoon, Z. Wang, and D. S. Park, "An available database for the research of finger vein recognition," in *Proc. 6th Int. Congr. Image Signal Process. (CISP)*, vol. 1, Dec. 2013, pp. 410–415.

[8] A. Kumar and Y. Zhou, "Human identification using finger images," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2228–2244, Apr. 2012.

[9] B. T. Ton and R. N. J. Veldhuis, "A high quality finger vascular pattern dataset collected using a custom designed capturing device," in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2013, pp. 1–5.

[10] H. Qin and M. A. El-Yacoubi, "Deep representation-based feature extraction and recovering for finger-vein verification," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 8, pp. 1816–1829, Aug. 2017.

[11] Y. Fang, Q. Wu, and W. Kang, "A novel finger vein verification system based on two-stream convolutional network learning," *Neurocomputing*, vol. 290, pp. 100–107, May 2018.

[12] W. Yang, C. Hui, Z. Chen, J.-H. Xue, and Q. Liao, "FV-GAN: Finger vein representation using generative adversarial networks," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 9, pp. 2512–2524, Sep. 2019.

[13] H. Qin and P. Wang, "Finger-vein verification based on LSTM recurrent neural networks," *Appl. Sci.*, vol. 9, no. 8, p. 1687, Apr. 2019.

[14] E. Jalilian and A. Uhl, "Enhanced segmentation-CNN based finger-vein recognition by joint training with automatically generated and manual labels," in *Proc. IEEE 5th Int. Conf. Identity, Secur., Behav. Anal. (ISBA)*, Jan. 2019, pp. 1–8.

[15] J. Zeng, F. Wang, J. Deng, C. Qin, Y. Zhai, J. Gan, and V. Piuri, "Finger vein verification algorithm based on fully convolutional neural network and conditional random field," *IEEE Access*, vol. 8, pp. 65402–65419, 2020.

[16] K. J. Noh, J. Choi, J. S. Hong, and K. R. Park, "Finger-vein recognition based on densely connected convolutional network using score-level fusion with shape and texture images," *IEEE Access*, vol. 8, pp. 96748–96766, 2020.

[17] X. Li, J. Lin, Y. Pang, L. Huang, L. Zhong, and Z. Li, "Fingertip blood collection point localization research based on infrared finger vein image segmentation," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[18] T. Wollmann, C. S. Eijkman, and K. Rohr, "Adversarial domain adaptation to improve automatic breast cancer grading in lymph nodes," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 582–585.

[19] M. Javanmardi and T. Tasdizen, "Domain adaptation for biomedical image segmentation using adversarial training," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 554–558.

[20] C. Chen, Q. Dou, H. Chen, J. Qin, and P. A. Heng, "Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation," in *Proc. AAAI Conf. Artif. Intell.*, 2019, vol. 33, no. 1, pp. 1–8.

[21] M. Wang, D. Zhang, J. Huang, P.-T. Yap, D. Shen, and M. Liu, "Identifying autism spectrum disorder with multi-site fMRI via low-rank domain adaptation," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 644–655, Mar. 2020.

[22] Q. Liu, Q. Dou, L. Yu, and P. A. Heng, "MS-Net: Multi-site network for improving prostate segmentation with heterogeneous MRI data," *IEEE Trans. Med. Imag.*, vol. 39, no. 9, pp. 2713–2724, Sep. 2020.

[23] S. Zhao, G. Wang, S. Zhang, Y. Gu, Y. Li, Z. Song, P. Xu, R. Hu, H. Chai, and K. Keutzer, "Multi-source distilling domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 7, pp. 1–9.

[24] Z. Zhou, L. Qi, X. Yang, D. Ni, and Y. Shi, "Generalizable cross-modality medical image segmentation via style augmentation and dual normalization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 20824–20833.

[25] K. He, W. Ji, T. Zhou, Z. Li, J. Huo, X. Zhang, Y. Gao, D. Shen, B. Zhang, and J. Zhang, "Cross-modality brain tumor segmentation via bidirectional global-to-local unsupervised domain adaptation," 2021, *arXiv:2105.07715*.

[26] X. Han, L. Qi, Q. Yu, Z. Zhou, Y. Zheng, Y. Shi, and Y. Gao, "Deep symmetric adaptation network for cross-modality medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 41, no. 1, pp. 121–132, Jan. 2022.

[27] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.

[28] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," 2016, *arXiv:1602.07360*.

[29] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

[31] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6848–6856.

[32] N. Ma, X. Zhang, H. T. Zheng, and J. Sun, "ShuffleNet v2: Practical guidelines for efficient CNN architecture design," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 116–131.

[33] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More features from cheap operations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1577–1586.

[34] Y. Li, Y. Chen, X. Dai, D. Chen, M. Liu, L. Yuan, Z. Liu, L. Zhang, and N. Vasconcelos, "MicroNet: Towards image recognition with extremely low FLOPs," 2020, *arXiv:2011.12289*.

[35] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, *arXiv:1503.02531*.

[36] N. Passalis and A. Tefas, "Learning deep representations with probabilistic knowledge transfer," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 268–284.

[37] X. Jin, B. Peng, Y. Wu, Y. Liu, J. Liu, D. Liang, J. Yan, and X. Hu, "Knowledge distillation via route constrained optimization," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1345–1354.

[38] S. Du, S. You, X. Li, J. Wu, F. Wang, C. Qian, and C. Zhang, "Agree to disagree: Adaptive ensemble knowledge distillation in gradient space," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 12345–12355.

[39] D. Chen, J. P. Mei, Y. Zhang, C. Wang, Z. Wang, Y. Feng, and C. Chen, "Cross-layer distillation with semantic calibration," in *Proc. AAAI Conf. Artif. Intell.*, 2021, vol. 35, no. 8, pp. 1–9.
[40] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, Munich, Germany. Springer, 2015.
[41] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539.

YINGXIN LI was born in 1975. He received the master's degree from Shanghai Maritime University, in 2007. He is currently a Senior Engineer. He is also engaged in medical image processing and medical information research with Jiangmen Central Hospital.

YUCONG CHEN received the M.S. degree in computer technology from Wuyi University, Guangdong, China, in 2023. He is currently with Jiangmen Central Hospital. His research interests include image segmentation and biometric recognition.

JUNYING ZENG (Member, IEEE) received the Ph.D. degree in physical electronics from Beijing University of Posts and Telecommunications, Beijing, China, in 2008, and the master's degree in physical electronic from Yunnan University, Yunnan, China, in 2005. Since June 2008, he has been with the Department of Intelligence Manufacturing, Wuyi University, Guangdong, China, where he is currently an Associate Professor. His research interests include image understanding, deep learning, and signal processing.

CHUANBO QIN received the B.S. and M.S. degrees from Wuyi University (WYU), China, in 2004 and 2008, respectively, and the Ph.D. degree from the Department of Automation Science and Engineering, South China University of Technology, in 2015. He is currently a Lecturer with WYU University. His main research interests include medical image segmentation and biometric recognition.

WENGUANG ZHANG was born in 1988. He received the master's degree in neurosurgery from China Medical University, in 2013. He is currently the Deputy Chief Physician in neurosurgery and the Sub-Specialist in functional neurosurgery with Jiangmen Central Hospital, Guangdong. His research interests include preoperative assessment and surgical treatment of traumatic brain injury, cerebrovascular disease, brain tumors, drug-refractory epilepsy, and Parkinson's disease.