Single Source Domain Generalization for Palm Biometrics

Congcong Jia[a], Xingbo Dong[a,*], Yen Lung Lai[a], Andrew Beng Jin Teoh[c], Ziyuan Yang[b], Xiaoyan Zhang[a], Liwen Wang[a], Zhe Jin[a], Lianqiang Yang[a]

[a]*Anhui Provincial Key Laboratory of Secure Artificial Intelligence, School of Artificial Intelligence, Anhui University, Hefei, China*
[b]*College of Computer Science, Sichuan University, Chengdu, China*
[c]*School of Electrical and Electronic Engineering, College of Engineering, Yonsei University, Seoul, Republic of Korea*

Abstract

In palmprint recognition, domain shifts caused by device differences and screening conditions pose significant challenges. Existing approaches often require multiple source domains for effective domain generalization (DG), limiting their applicability in single-source domain scenarios.To address this challenge, we propose PalmRSS, a novel Palm Recognition approach on top of Single Source Domain Generalization (SSDG). PalmRSS reframes the SSDG problem as a DG problem by partitioning the source domain dataset into subsets and employing image alignment and adversarial training. PalmRSS exchanges low-level frequencies of palm data and performs histogram matching between samples to align spectral characteristics and pixel intensity distributions. Experiments demonstrate that PalmRSS outperforms state-of-the-art methods, highlighting its effectiveness in single source domain generalization.

Keywords: Palmprint recognition, Single source domain generalization, Open-set recognition, Low-level frequencies, Histogram matching

1. Introduction

The palm contains unique and stable features, such as palm lines, creases, and patterns, which are highly distinctive and can be reliably captured and analyzed. These

*Corresponding author

traits make palmprint recognition robust and accurate biometrics, even in challenging environments. Moreover, palmprint recognition allows for contactless capture of palm images. This implies that individuals can be identified without needing physical contact, enhancing convenience and hygiene, especially when contact may be undesirable or unavailable. Consequently, palm-based modalities such as palmprint and palm vein [1] stand out amidst many biometric methods.

Palmprint recognition systems extract unique features from palmprint images and compare them to stored templates in a database, with various feature extraction methods having been proposed. They include line-based [2], local direction encoding-based [3], texture descriptor-based methods [4] and learning-based [1]. With the advent of deep learning, deep learning-based methodologies have become prominent in palmprint and palm vein recognition, markedly enhancing performance [5]. As a result, integrating learning-based palmprint and palm vein recognition systems into everyday devices like smartphones has grown increasingly common. The cross-device utilization of palm biometrics promises a seamless and secure authentication experience, surpassing traditional passwords or PINs. Nevertheless, the heterogeneous visual characteristics among images across devices present a substantial challenge. Existing methods, such as CO3Net [6] and CCNet [5], often fail to address domain shift challenges, leading to subpar performance in cross-device recognition scenarios.

As illustrated in Fig. 1, domain shifts occur due to device differences and screening conditions. The disparity between different domains (devices or datasets) manifests as variations in brightness and gain-like conditions of palm images. These variations in the source and target domains primarily manifest as pixel-wise transformations. Additionally, we observed distinct differences in the amplitude component of the Fourier spectrum after the Fourier Transform, which encapsulates low-level statistics of the original signal across different devices. In response to this domain shifting challenge, cross-dataset palmprint recognition [7] with integrated domain generalization (DG) [8]

emerges as a valuable approach. The DG method aims to learn a model capable of achieving good performance on unseen domains, offering a promising solution to the quest for robust and adaptable palmprint recognition systems.

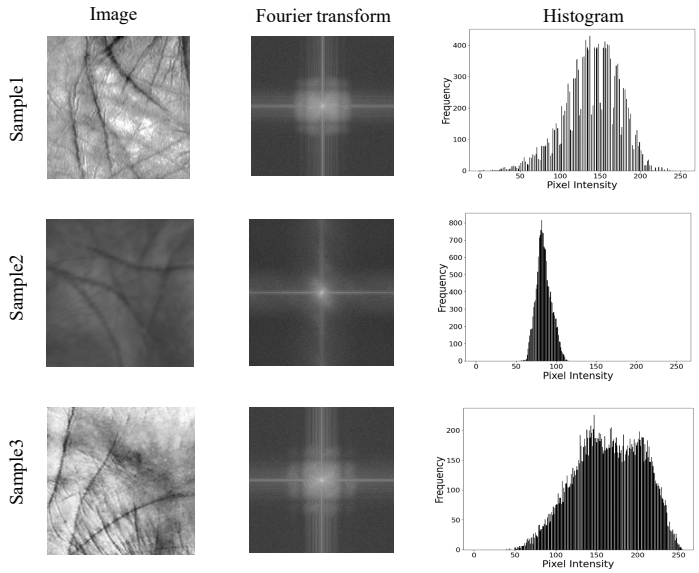

Figure 1: Sample 1, Sample 2, and Sample 3 correspond to different palmprint images of the three datasets; from top to bottom are the dataset of the PolyU [9], Tongji [10], and IITD [11] respectively. The second column is the Fourier transform of the sample, and the third column is the histogram of the sample. Each histogram's horizontal axis represents "Pixel Intensity" and the vertical axis represents "Frequency".

For instance, in [12], the authors propose a distribution-based loss to address the domain shift issue caused by heterogeneous image characteristics. This enhancement aims to improve the model's cross-device recognition performance by magnifying the disparity between intra-class and inter-class similarities. The final loss function comprises the positive histogram loss to attract positive pairs, the negative histogram loss to repel negative pairs, and the ArcFace loss for classification. However, in this study, achieving positive pairs requires identical identities to be present in two datasets collected using different devices. In practice, this is often challenging to accomplish.

Another recent work by Shao et al. [13]. Namely, Palmprint Data and Feature Generation (PDFG) has delved into a setting of cross-dataset palmprint recognition with the

unseen target dataset. The target dataset's palmprint images are completely unavailable during training in this setup. Consequently, the model is tasked with achieving robust generalization to unseen target datasets. More specifically, in scenarios where there exist N source datasets, denoted as $D^1, \ldots, D^N$, a Fourier-based data augmentation method is firstly applied to generate more training data to encourage the model to see more image distributions. Then, consistent loss and dataset-aware triplet loss are utilized to extract features adaptive to the source and augmented datasets, which are further generalized to unknown target datasets. [13] essentially involves the exchange of feature information from palm images across at least two distinct datasets (i.e., $D^1, D^2 \ldots, N > 1$) during model training. The trained model is then expected to demonstrate robust generalization when applied to unseen data from another target set D^t.

In this perspective, it can be argued that the model's generalization capability may rely on the correlation between additional source datasets (e.g., D^2) and the target set D^t, thus suggesting a dependency of domain generalization on D^2. This raises a compelling question: Can a model effectively generalize from a single source domain to numerous unseen target domains? In other words:

"How can we optimize the model's generalization performance when training is limited to just one ($N = 1$) domain?"

The above question is intricately tied to a fundamental pursuit within machine learning research, termed "Single Source Domain Generalization" (SSDG) [14]. In essence, SSDG seeks to equip models with the ability to generalize effectively to unseen domains, even when trained solely on data from a single source. As depicted in Fig. 2 (case (d) and (e)), it is worth noting that SSDG can be interpreted as zero-shot learning [15], which often involves open-set recognition [16] , where the model must identify instances from classes not seen during training. In light of this interpretation, our approach to tackling the SSDG problem can be easily adapted for the open-set

palmprint recognition problem.

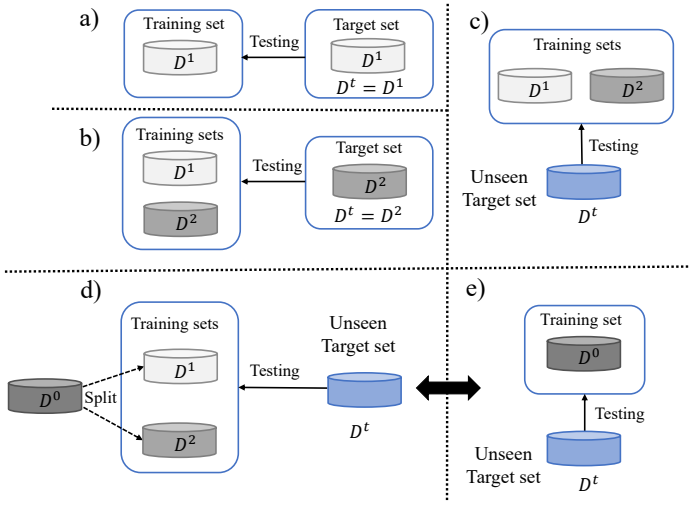

Figure 2: Illustrations of a) Traditional palmprint recognition, where the training and test sets originate from the same dataset. b) Traditional cross-dataset palmprint recognition, where the test set is incorporated into the training set. [7]. c) Cross-dataset palmprint recognition with target dataset, where the test target set is unseen during the training phase [13]. d) Our case, where the training sets are treated as subsets of a single source dataset D^0, and the test target set remains unseen during the training phase. e) An open-set scenario [17] , the training set does not include any samples belonging to the unseen target set, which is equivalent to our case in d) [1].

Despite the wealth of research on SSDG across various fields such as medical imaging segmentation [18], natural image classification [14], hyperspectral image classification [19], To our best knowledge, there is a noticeable gap in research concerning SSDG for palmprint recognition as of our investigations.

In this paper, we proposed a new **Palm R**ecognition scheme on top of **S**ingle **S**ource domain generalization, namely PalmRSS. By dividing the source domain dataset (D^0) into subsets (D^1 and D^2), PalmRSS reframes the SSDG problem as a DG problem. Building upon the generated subsets D^1 and D^2, inspired by the Fourier Domain Adaptation (FDA) [20], we propose image alignment across samples in the source dataset and employ adversarial training to optimize our model. Initially, we exchange low-level

[1]Here, we considered the training process to encompass supervised, unsupervised, or a combination of both training approaches.

frequencies of palm data from different users; this exercise aids in aligning the spectral characteristics across various samples. Additionally, we perform direct histogram matching between users' palm images, aligning pixel intensity distribution across different samples.

Despite its simplicity, our experimental results demonstrate recognition performance that outperforms the state-of-the-art methods under an SSDG framework. While we do not interpret our finding as the optimal way to achieve domain generalization for palmprint recognition, our stance acknowledges that previous methods could be less effective in achieving the same goal, where they require more than a single source domain dataset for training [13], or adopting sophisticated domain expansion strategies tailored for domain generalization. Our results demonstrate better achievable performance than the methods above in subsequent experiments, as outlined in Section 4.

In summary, our main contributions are as follows:

1. We introduce PalmRSS, a novel approach to palm biometric recognition within the SSDG framework. This framework breaks new ground in the domain of palmprint recognition by addressing the challenge of generalizing effectively to unseen domains, even with limited training data from a single source.
2. We propose two techniques for data alignment across subsets within the source domain (D^0). Specifically, we employ Fourier alignment transform and histogram matching to mitigate distribution differences between subsets. These techniques align palm data across samples, reducing domain shift and improving model generalization.
3. To ensure the effectiveness of the alignment transformations, we introduce domain adversarial loss and feature similarity loss. These losses regularize the alignment process, guiding the model to learn robust representations of domain variations while preserving discriminative features crucial for recognition tasks.
4. We conduct extensive experiments to validate the efficacy of PalmRSS. Our ex-

periments encompass cross-dataset and single-dataset evaluations, comparing PalmRSS against state-of-the-art methods. The results demonstrate the superior performance of PalmRSS across various experimental setups, underscoring its efficacy in addressing the SSDG challenge in palmprint recognition.

2. Related Work

2.1. Palmprint Recognition

Handcrafted Methods: One of the pioneering works in this field was PalmCode by Zhang et al. [3]. PalmCode was inspired by the success of iris feature extraction, specifically the IrisCode [21], which utilized a 2D Gabor phase coding scheme. Following Zhang et al.'s work, Kong et al. [22] introduced a more robust version of PalmCode known as Competitive Code (CompCode). Expanding on the success of CompCode, researchers have delved into new coding principles to enhance competition mechanisms, thereby fostering the advancement of palmprint recognition technology.

Building on the success of coding-based palmprint recognition [3], Gabor filters have been widely adopted for palmprint feature encoding. Along this line of research, Guo et al. [23] proposed a new palmprint feature extraction method, namely binary orientation co-occurrence vector (BOCV), which is more robust to slight image rotation compared to traditional competitive codes.

Yang et al. [24] introduced a Multiple-order Texture Co-occurrence Code (MTCC), combining first-order and second-order texture features. The second-order texture is derived from the convolution of one-dimensional texture with one-dimensional Gabor filters. Gabor feature encoding is utilized to generate both the first-order Texture Co-occurrence Code (1TCC) and the second-order Texture Co-occurrence Code (2TCC) from their respective textures, which are then fused at the score level to produce MTCC.

Learning-Based Methods: In addition to the previously outlined methods, which rely on human-designed filters for the extraction of palmprint features, deep learning-

based approaches have been extensively embraced and have demonstrated significant efficacy. Genovese et al. [25] proposed a network called PalmNet. PalmNet introduces a CNN architecture that utilizes unsupervised Gabor response and principal component analysis (PCA) to optimize palmprint-specific filters, achieving high recognition accuracy without class labels during training. Selma and Djamel et al. proposed a simplified Palmnet-Gabor to improve Palmnet for fast palmprint recognition [26]. The method utilizes Log-Gabor filters to increase the contrast of palmprint features in preprocessing. Then, feature selection and dimensionality reduction are used to reduce the number of features, ensure decent recognition accuracy, and reduce the complexity of the deep learning model.

Hybrid Methods: The hybrid approach combines learning and handcrafted algorithms to achieve the task. CCNet [5] is one of the latest representative works with state-of-the-art performance. The primary component of CCNet is inherited from CompNet [27], and the competition mechanism is expanded based on CO3Net [6]. Specifically, it involves the utilization of Competitive Blocks(Comp Block, CB), which employ 2-D Gabor phase encoding to extract competitive features from palm images. The Gabor filter, a fundamental component of this process, is defined as follows:

$$G(x, y; \sigma, \gamma, \mu, \psi, \theta) = e^{-\frac{\gamma^2 x'^2 + y'^2}{2\left(2\sqrt{2}\sigma\right)^2}} \cos(2\pi\mu x' + \psi), \tag{1}$$

where the coordinates (x,y) represent the index within the filter. The orientation of the Gabor filter is denoted as θ, the standard deviation of the Gaussian envelope as σ, and the sine wave frequency as μ. Additionally, γ and ψ correspond to the Gaussian function's spatial aspect ratio and scale factor.

When given an input palmprints data F_{in} to the Comp Block, Soft Competition Code (SCC) [27] is employed with softmax function (denoted as Softmax) to extract the ranking relationship along channel dimensions (c) and spatial dimensions (x-axis and y-axis), formulated as follows:

$$
\begin{aligned}
F_c &= \text{Softmax}_c(F_{\text{in}}) \\
F_x &= \text{Softmax}_x(F_{\text{in}}) \\
F_y &= \text{Softmax}_y(F_{\text{in}}).
\end{aligned} \tag{2}
$$

Here, F_c, F_x, and F_y represent the results of applying the softmax function along the channel (c), x-axis, and y-axis dimensions of the input feature F_{in}, respectively.

CCNet employs three parallel texture extraction branches to extract competition features at multiple scales. As shown in Fig. 3, this is achieved by utilizing Gabor filters of varying sizes: 7×7, 17×17, and 35×35, corresponding to large, middle, and small-scale output features. Each branch consists of two learnable Gabor layers and two Comp Blocks, joined alternately to extract multi-order texture features. Learnable Gabor filters are adopted to optimize these hyper-parameters.

To reduce the feature size while retaining channel and spatial competition results, a weighted sum operation is introduced, formulated as:

$$
F_{\text{out}} = w_z \times F_c + w_s \times (F_x + F_y), \tag{3}
$$

where F_{out} is the output of the Comp Block, w_z is the weight of the channel competition, and w_s is the weight of the spatial competition.

2.2. Cross-dataset Palmprint Recognition

Significant contributions have been made in Cross-dataset Palmprint Recognition. For instance, Du et al. [28] introduced a regularized adversarial domain adaptive hashing (R-ADAH) framework for palmprint recognition. The R-ADAH framework exhibits strong generalization capabilities and does not require matching identities between source and target domains. Additionally, it has potential applications across various biometric systems to address cross-domain identification challenges.

Shao et al. [29] introduced the Joint Pixel and Feature Alignment (JPFA) for en-

hancing palmprint recognition across diverse datasets. This approach employs a two-stage alignment process to obtain adaptive features. Initially, the source images transform pseudo-images through a deep transfer model, thereby minimizing the disparities between datasets. Subsequently, the target image-source image and target image-pseudo-image comparisons enable extracting essential adaptive features, achieving pixel-level alignment and feature alignment, respectively.

Recognizing that palmprint features vary across different spectra, with varying emphasis even for the same individual, Fei et al. [7] proposed a cross-spectral PalmGAN. This work incorporates constraints into the learning objective function, such as the feature variance maximum and the feature gap minimization constraints, enabling the effective learning of discriminative spectrum-invariant features from multispectral palmprint images.

Shao and Zhong [30] proposed a multi-objective cross-dataset palmprint recognition method, termed multi-teacher distillation (DFMT), which leverages knowledge distillation and domain adaptation. An independent teacher extractor is constructed for each target dataset in their approach. The source dataset is paired with each target dataset to facilitate the learning of these teacher extractors. However, a significant limitation of this method is its reliance on multiple datasets; both the source and multiple target datasets are required during the training process. Although this approach effectively addresses the issue of handling multiple target datasets, it does not account for scenarios where the target datasets are unseen by the training model.

2.3. Motivation of PalmRSS

In this work, PalmRSS partitions the source domain dataset (D^0) into two subsets (D^1 and D^2). We introduce image alignment across samples within the source dataset and utilize adversarial training to optimize our model without needing any target dataset (D^t) samples. The novel approach of dividing the source domain dataset and applying histogram matching and Fourier transform between the subsets is an uncon-

ventional yet rational method for domain adaptation. The motivations for our approach are as follows:

- Simulating Domain Shift and Relax to the DG task: By dividing the source domain into two subsets and applying transformations between them, we can simulate a domain shift within the source domain. This simulated shift prepares the model to handle variations and discrepancies that it might encounter in a real target domain, enhancing its robustness.

- Ensuring Diversity: The source domain may encompass a variety of samples. By partitioning this domain and applying transformations inspired by FDA [20], we can expose the model to a diverse array of examples. These transformations promote the learning of features invariant to variations in intensity and frequency distributions, thereby enhancing the model's ability to acquire general and robust features. This process reduces sensitivity to domain-specific variations, which is essential for successful domain adaptation.

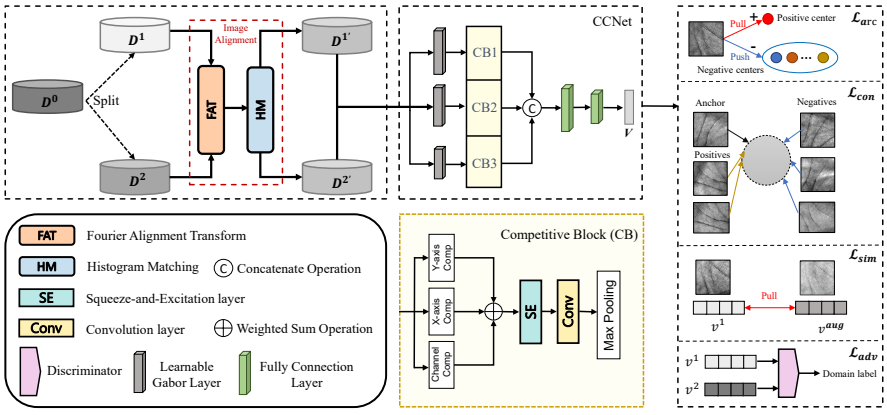

Figure 3: The proposed model consists of three main components: image alignment module, CCNet module, and loss functions module. D^0, D^1 and D^2 correspond to the representations in Fig. 2. d) and e), respectively. The image alignment module consists of FAT and HM. The transformed image is sent to CCNet to get the corresponding feature vector V. The model is optimized by the loss functions.

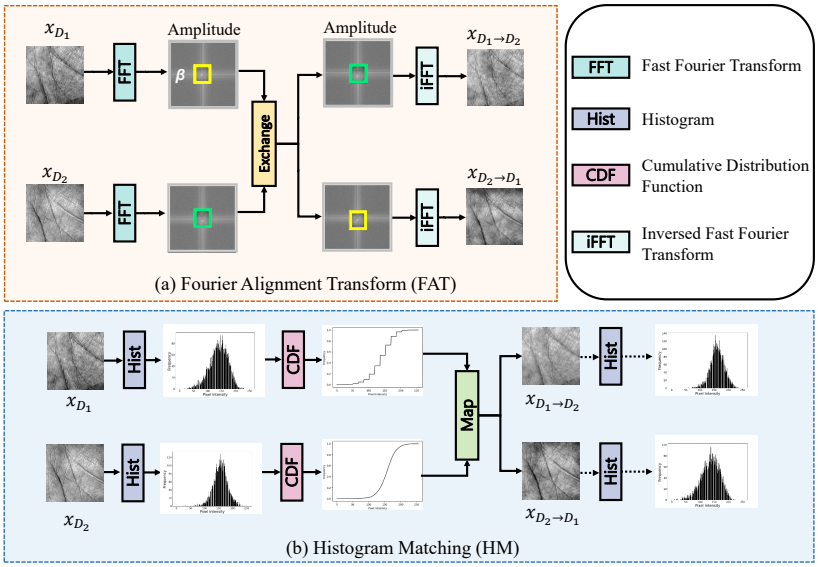

Figure 4: (a) Illustrates the Fourier Alignment Transform (FAT). The image x_{D^1} undergoes a Fourier transform, and a selected portion of its amplitude spectrum, determined by the parameter β, is exchanged with the corresponding portion from x_{D^2}. This modified amplitude spectrum is then used to generate a new image via the inverse Fourier transform. (b) Depicts the Histogram Matching (HM) process using x_{D^1} as an example. The "Map" indicates that the histogram distribution of x_{D^1} is adjusted to match that of x_{D^2}, resulting in a new image $x_{D^1 \to D^2}$.

3. Methodology

3.1. Overview

As illustrated in Fig. 3, on top of the data splitting strategy, the overall process comprises three main components: the image alignment module, the CCNet module, and the loss function module. Firstly, we partition the source domain dataset D^0 into subsets D^1 and D^2. Secondly, for any image from D^1 and D^2, we perform a Fourier alignment transform and histogram matching. Subsequently, the transformed images are fed into CCNet to obtain the corresponding feature vector. Finally, a combination of recognition and domain adversarial loss is used to optimize the model. The specific details of the method are described as follows.

3.2. Image Alignment

3.2.1. Fourier Alignment Transform

As depicted in Fig. 1, palmprint images obtained from different devices exhibit noticeable disparities after the Fourier transform. Additionally, inspired by FDA [20], we mitigate the distinctions between dataset distributions by exchanging the low-frequency spectrum of different datasets, such as D^1 and D^2. Considering that the primary objective of biometric palmprint recognition is to furnish convenient and dependable recognition performance suitable for real-time recognition tasks, our approach prioritizes minimizing model complexity. Hence, we employ an efficient Fourier alignment transform (FAT) to achieve straightforward alignment of low-level statistics between source and target distributions.

More specifically, this entails computing the Fast Fourier Transform (FFT) of each input image and subsequently substituting the low-level frequencies of the target images with those from the source images before reconstructing the image for training using the inverse FFT (iFFT). However, although the original FDA [20] is proposed for image segmentation, it does not explicitly tackle the more challenging problem, particularly considering the constraint of having only a single source domain available for training.

FAT tackles this challenge by partitioning a single source dataset, denoted as D^0, into two subsets, D^1 and D^2. This strategy aligns with the PDFG[13], where D^1 and D^2 serve as the model's training set. Unlike D^1 and D^2 from [13], which are derived from different sources, our partitioned subsets originate from a single source. The primary objective is to train a model capable of generalizing its learned features to an unseen target set D^t (target domain). With this alignment, we can effectively operate within the SSDG framework utilizing only a single source dataset, D^0, for palmprint recognition. We denote a palmprint image from D^0 as $x \in \mathbb{R}^{H\times W\times 3}$, undergoing Fourier transform

as:

$$\mathcal{F}(x)(m,n) = \sum_{h=1}^{H}\sum_{w=1}^{W} x(h,w)e^{-j2\pi\left(\frac{h}{H}m+\frac{w}{W}n\right)}, \quad j^2 = -1. \tag{4}$$

Here, $\mathcal{F}(x)(m,n)$ is the transformed output, $x(h,w)$ is the pixel intensity, and m and n are coordinates in the frequency domain.

Meanwhile, a mask M_β, whose value is zero except the center region, is defined as

$$M_\beta(h,w) = \mathbb{1}_{(h,w)\in[-\beta H:\beta H,-\beta W:\beta W]}, \tag{5}$$

where $\beta \in (0,1)$, and in our case, $\beta = 0.1$.

After splitting D^0 into $D^1 \subset D^0$ and $D^2 \subset D^0$, let $x_{D^1} \sim D^1$ and $x_{D^2} \sim D^2$ denote randomly sampled palmprint images. The FAT can be described as follows:

$$x_{D^1\to D^2} = \mathcal{F}^{-1}([M_\beta \circ \mathcal{F}^A(x_{D^2}) + (1-M_\beta)\circ\mathcal{F}^A(x_{D^1}), F^P(x_{D^1})]), \tag{6}$$

where $\mathcal{F}^A$ and $\mathcal{F}^P$ denote the amplitude and phase components of $\mathcal{F}(.)$ respectively. $\mathcal{F}^{-1}$ represents the inverse of $\mathcal{F}(.)$, also known as the inverse Fast Fourier Transform (iFFT).

This process involves replacing the low-frequency components ($\mathcal{F}^A(x_{D^1})$) of the amplitude of x_{D^1} with those from another image x_{D^2}, while keeping the phase component ($\mathcal{F}^P(x_{D^1})$) of x_{D^1} unchanged. The modified spectral representation is then returned to the image domain, resulting in a transformed image denoted as $x_{D^1\to D^2}$. The overall process is illustrated in Fig. 4.

Although FAT resembles FDA [20], in our approach, the datasets D^1 and D^2 are exclusively derived from a single source dataset D^0. As highlighted by [20], aligning the low-level statistics between source and target distributions can enhance performance in unsupervised domain adaptation. Accordingly, FAT confers an additional "zero-shot" capability to FDA. This capability allows our method to be effectively applied to SSDG in palmprint images, enabling it to handle target domains not encountered during the

training phase.

Importantly, FAT minimizes the dependency of domain generalization on the available source dataset and the associated training cost, particularly for small-size source datasets such as D^0. We address the question posed in Section 1 by optimizing the model's generalization performance when training is limited to just one domain.

3.2.2. Histogram Matching

From Fig. 1, it's evident that palmprint images collected by the different devices exhibit notable differences in their histograms across different spectra. Hence, drawing from the FDA's transformation, we endeavored to mitigate disparities in distribution among various subsets using histogram matching.

Histogram matching is commonly employed to align pixel-level distributions, facilitating image alignment based on low-level statistics in the spatial domain. In unsupervised domain adaptation, this alignment can make the statistical properties of images in the target domain more akin to those in the source domain. Such alignment helps reduce the domain gap and enhances the performance of the adapted model. In our SSDG setting, we utilize histogram matching to increase statistical diversity in the spatial domain, thereby improving the model's generalization capabilities.

Histogram matching is achieved by computing the cumulative distribution functions (CDFs) of pixel (intensity) values up to i for x_{D^1} and x_{D^2}:

$$\begin{aligned} C_{D^1}(i) &= \sum_{j=0}^{i} h_{D^1}(j), \\ C_{D^2}(i) &= \sum_{j=0}^{i} h_{D^2}(j), \end{aligned} \tag{7}$$

where ($C_{D^1}(i)$,$C_{D^2}(i)$) and ($h_{D^1}(j)$, $h_{D^2}(j)$) are the CDFs and histograms of x_{D^1} and x_{D^2}, respectively. A mapping function $T(.)$, described as

$$T(i) = C_{D^2}^{-1}(C_{D^1}(i)), \tag{8}$$

where $C_{D^2}^{-1}$ denotes the inverse function of C_{D^2}, can be used to map individual intensity level of x_{D^2} to x_{D^1}, resulting the transformed image:

$$x_{D^1 \to D^2}(h, w) = T(x_{D^1}(h, w)). \tag{9}$$

The overall process is illustrated in Fig. 4.

3.3. Loss Functions

We utilized CCNet as our backbone and trained it on datasets D^1 and D^2 to develop our final model. During training, we incorporated a hybrid loss function to facilitate recognition and a domain adversarial loss to enhance domain generalization.

3.3.1. Hybrid Recognition Loss for D^1:

We adopt the hybrid loss from CCNet and introduce a feature similarity loss to minimize intra-class variations in the feature space. Specifically, we employ the ArcFace loss [31], supervised contrastive learning loss [32], and feature similarity loss for supervised training on subset D^1.

More precisely, the ArcFace loss is defined as:

$$L_{arc} = -\frac{1}{N} \sum_{i=1}^{N} \log \left(\frac{e^{s(\cos(\theta_{yi}+m))}}{e^{s(\cos(\theta_{yi}+m))} + \sum_{j=1, j\neq y_i}^{J} e^{s \cos \theta_j}} \right) \tag{10}$$

where N and J denote the numbers of samples and classes. θ_j is the angle between the weight W_j and the embedding v_i, m is the additive angular margin penalty, s is the scale factor.

The supervised contrastive loss is given as follows:

$$\mathcal{L}_{con} = -\sum_{i \in I} \frac{1}{|p(i)|} \sum_{p \in P(i)} \log \frac{exp(z_i \cdot z_p / \tau)}{\sum_{a \in A(i)} exp(z_i \cdot z_a / \tau)}, \tag{11}$$

where

- $I \equiv \{1, ..., 2N\}$ denotes the batch's set of contrastive sample pairs.
- $I \equiv I \setminus \{i\}$ represents the positive sample in index i.
- $P(i) \equiv \{p \in A(i) : y_i = y_p\}$ indicates the index set of forward samples different from the index i, where y_i represents the label of the i-th sample in the batch.
- $|P(i)|$ signifies the number of samples in $P(i)$.
- z_p and z_a are the positive and anchor embeddings, respectively.
- τ is the temperature hyperparameter.

Further minimizing the intra-class distance in feature space can enhance the model's generalization performance. Therefore, we propose a feature similarity loss to encourage the model to learn more discriminative features that exhibit resilience to the higher variation. The feature similarity loss is given as:

$$L_{sim} = \frac{1}{N} \sum_{i=1}^{N} d(v^1, v^1_{aug}) \tag{12}$$

where $d(\cdot)$ is the cosine distance, N denotes the numbers of samples, and v^1_{aug} denotes the augmented instances from the original input data.

The final hybrid loss function is defined as follows:

$$\mathcal{L}_{hyb} = w_{arc} \times \mathcal{L}_{arc} + w_{con} \times \mathcal{L}_{con} + \lambda \mathcal{L}_{sim}, \tag{13}$$

where w_{arc} , w_{con} , and λ are the weights values, in which we set to 0.8, 0.1 and 0.1 for the cross entropy loss $\mathcal{L}_{arc}$, supervised contrastive loss $\mathcal{L}_{con}$ and feature similarity loss $\mathcal{L}_{sim}$, respectively.

3.3.2. Domain Adversarial Loss:

Domain adversarial loss is commonly employed in domain adaptation problems to address the distribution mismatch between a labeled source domain and an unlabeled

target domain [33]. The embedding learned becomes domain-invariant by training a domain classifier to distinguish between the domains and simultaneously optimizing the task classifier to perform well on the source domain. This process aligns the domains, thereby improving model generalization across different distributions and enhancing performance on the target domain.

The domain adversarial loss is defined as:

$$\mathcal{L}_{adv} = -\mathbb{E}_{v_1 \sim \mathcal{D}^1}[\log(A(v_1))] - \mathbb{E}_{v_2 \sim \mathcal{D}^2}[\log(1 - A(v_2))], \tag{14}$$

where $A(\cdot)$ is the domain classifier that predicts whether a given feature representation belongs to the source or target domain. v_1 and v_2 are the embedding extracted from D^1 and D^2 subsets, respectively.

Summing up, the final loss function in our training process can be described as follows:

$$\mathcal{L} = \mathcal{L}_{adv} + \lambda \mathcal{L}_{hyb}. \tag{15}$$

4. Experiments and Results

4.1. Palmprint and Palmvein Datasets

PolyU [9] PolyU collected palmprint images from 193 individuals, 86% of whom were under the age of 30, using a contact-based palmprint acquisition device. The data was collected twice, approximately two months apart. Each subject provided approximately 20 images of the palm at a time, resulting in about 40 images per person. The database contains 7,752 images from 193 individuals, with 7,560 selected as valid images.

Tongji [10] Tongji collected palmprint images of 300 individuals using a contactless palmprint collection device, with 87% of the subjects under the age of 30. The data

was collected twice, about two months apart. Subjects provided images of 20 palms at a time, resulting in 12,000 palmprint images from 600 different palms.

IITD [11] IITD collected palmprint images of 460 people using a contactless palmprint acquisition device. Subjects provided about five images each time, resulting in 2,601 images. After preliminary processing, 2,300 images were selected as experimental data.

CasiaM [34] CasiaM collected 7,200 palmprint images from 100 individuals using a multispectral imaging contactless palmprint acquisition device. Subjects provided approximately six samples at a time, collected twice, over a month apart. In each acquisition, each sample contained six images of the palm, captured simultaneously using six different electromagnetic spectra corresponding to illuminator wavelengths of 460nm, 630nm, 700nm, 850nm, 940nm, and white light. For experimental data, images with wavelengths of 460nm, 700nm, and 850nm were selected, resulting in 3,600 images in the selected database.

Multi-Spectral (MS) [3] Multispectral collected palm images from 500 people using a contact palmprint collection device. Subjects provided approximately six samples at a time, collected twice, approximately two months apart. In each acquisition, each sample contained four images of the palm, simultaneously captured using four different spectra: red, green, blue, and near-infrared. The database contains a total of 24,000 images from 500 individuals.

4.2. Experiment Settings

We employed the Equal Error Ratio (EER) as a validation metric to test the performance. EER represents the point on the ROC curve where the False Acceptance Rate (FAR) equals the False Rejection Rate (FRR), with FRR calculated as one minus the Genuine Acceptance Rate (GAR), i.e., FRR = 1-GAR. EER is inversely proportional to performance, meaning lower EER values indicate higher performance.

For loss function Eq. 15, we set λ to 1. We essentially target the single-source

domain generalization recognition task. Therefore, three experiment settings, including two cross-domain recognition and one single-domain recognition, are adopted:

1. Cross-spectral: evaluations are conducted on palm datasets collected by the device under various spectral conditions. One spectrum is designated as the source domain, while the remaining spectrum serves as the target domain for completing the cross-domain experiment. The corresponding D^0 and D^t of cross-spectral are tabulated in Tab. 1;
2. Cross-device: evaluations are conducted using palmprint datasets collected by different devices (i.e., different datasets). Three palmprint datasets gathered from distinct devices and environments are chosen, with one serving as the source domain and the others as target domain datasets for cross-domain experiments. The corresponding D^0 and D^t of cross-device are tabulated in Tab. 1;
3. Single dataset: evaluations are conducted using a single palmprint dataset, allowing comparisons with existing studies.

Table 1: SSDG Experiment Settings.

Setting \ Dataset	D^0 ($D^1 \cup D^2$)	D^t
Cross-spectral	Red	Green, Blue, NIR
	Green	Red, Blue, NIR
	Blue	Red, Green, NIR
	NIR	Red, Green, Blue
	460	700, 850
	700	460, 850
	850	460, 700
Cross-device	Tongji	Red, Green, Blue, NIR, PolyU, IITD
	PolyU	Tongji, IITD
	IITD	Tongji, PolyU

4.3. Performance under Cross-domain Setting

Cross-spectral: Firstly, we conducted cross-spectral experiments by selecting different spectral datasets, designating one as D^0 and the other datasets as D^t. Tab. 2

presents the results of cross-spectral experiments on the MS dataset, where different spectral datasets are used as the source dataset (D^0) and target datasets (D^t). CCNet, CO3Net, and JPFA are included in our comparison as they focus on cross-dataset palmprint recognition, similar to our task. Both CCNet and CO3Net adopt the competitive mechanism, which is the backbone of our work.

Based on Tab. 2 we can observe:

- The EERs are generally lower when transitioning between visible spectrum bands (Red, Green, Blue). All methods perform better in these transitions compared to those involving NIR.
- Transitions involving the NIR spectrum generally result in higher EERs. This is likely due to the significant spectral difference (domain gap) between the visible spectrum and NIR.
- Ours method consistently outperforms the other existing works (CCNet, CO3Net, JPFA$^{\mathbf{Op}}$) across all spectral pairs, demonstrating significantly lower error rates. This indicates robustness and better performance in handling cross-spectral variability.

Tab. 3 presents results from cross-spectral experiments on the CasiaM dataset, with three different spectral bands: 460 nm, 700 nm, and 850 nm. Similarly, the results from Tab. 3 highlight that our method consistently outperforms CCNet, CO3Net, and DFMT across all spectral transitions, demonstrating its robustness and superior performance in cross-spectral experiments on the CasiaM dataset.

We also compare with existing domain adaption methods such as [13]. Tab. 4 presents the performance comparison results under a cross-spectral setting on the CasiaM dataset. The best performance in terms of EER and accuracy across different settings is highlighted in bold. We can observe From the results that our method demonstrates competitive performance, often achieving lower EERs and higher accuracies than the

other methods. Additionally, our method is evaluated using a single-source dataset under an SSDG setting, other methods are evaluated using multiple source datasets.

Cross-devices: Secondly, we mainly conducted cross-devices comparison experiments using the Tongji, PolyU, MS, and IITD datasets, following the SSDG scenario where each dataset served as D^0. In contrast, the others acted as D^t and were not involved during training. Tab. 5 presents the EER results of cross-dataset experiments. Based on the result, we can observe that our method consistently outperforms CCNet, CO3Net, and R-ADAH across all cross-dataset settings despite its simplicity and focus on a single source domain. The EERs for our method are the lowest in each case, demonstrating superior robustness and effectiveness in cross-dataset experiments.

Table 2: EERs of Cross-spectral setting on MS dataset.

Dataset (D^0 / D^t)	Red			Green			Blue			NIR		
Method	Green	Blue	NIR	Red	Blue	NIR	Red	Green	NIR	Red	Green	Blue
CCNet[5]	0.0278%	0.0611%	0.0091%	0.0026%	0%	0.2833%	0.0130%	0%	0.3000%	0.0167%	0.5889%	0.6140%
CO3Net[6]	0.1270%	0.1560%	0.0220%	0.0610%	0.0100%	0.6570%	0.1300%	0.0080%	0.8110%	0.0830%	1.5390%	1.3830%
JPFA$^{\mathbf{Op}}$[29]	0.0340%	0.0730%	0.2500%	-	-	-	0%	0.0550%	1.7700%	-	-	-
Ours	**0.0181%**	**0.0222%**	**0.0056%**	**0.0008%**	**0%**	**0.2500%**	**0.0107%**	**0%**	**0.2722%**	**0.0167%**	**0.0111%**	**0.0500%**

Op stands for the open set, " - " means no data from the original paper.

Table 3: EERs of Cross-spectral setting on CasiaM dataset.

Dataset(D^0 / D^t)	460		700		850	
Method	700	850	460	850	460	700
CCNet[5]	6.70%	3.67%	2.94%	1.39%	4.53%	2.33%
CO3Net[6]	10.72%	4.82%	3.56%	1.50%	5.83%	3.00%
DFMT[30]	10.21% *		-	-	-	-
Ours	**6.16%**	**3.33%**	**1.56%**	**1.28%**	**3.89%**	**2.11%**

" * " indicates that the experimental result is obtained with D^0 = 460 and D^t is the merge of 700 and 850.

Table 4: Performance Comparison with existing DA methods under Cross-spectral setting on CasiaM dataset.

	Source (D^0)	Target (D^t)	EER	Accuracy
PDFG[13]	460,630	700	3.79%	**97.67%**
	630,850		10.00%	**97.67%**
Ours	460		6.16%	92.50%
	850		**2.11%**	96.67%
PDFG[13]	630,850	460	10.97%	67.00%
Ours	700		**1.56%**	**98.33%**
	850		3.89%	94.17%
ADDA[35]	460,630	850	10.73%	78.00%
DIFEX[36]			12.02%	83.00%
R-ADAH[28]			11.60%	68.33%
PDFG[13]			3.72%	93.33%
PDFG[13]	460,700		1.48%	96.67%
Ours	460		3.33%	93.67%
	700		**1.28%**	**97.67%**

All experimental results except **Ours** are from [13].

Table 5: EERs of Cross-devices setting.

Dataset(D^0 / D^t)	Tongji						PolyU		IITD	
Method	Red	Green	Blue	NIR	PolyU	IITD	Tongji	IITD	PolyU	Tongji
CCNet[5]	0.811%	1.402%	1.017%	0.750%	1.570%	6.120%	2.220%	5.730%	2.010%	3.220%
CO3Net[6]	0.878%	1.580%	1.570%	0.783%	1.660%	7.640%	3.580%	8.740%	2.260%	3.720%
R-ADAH[28]	7.000%	6.200%	6.300%	7.700%	-	-	-	-	-	-
Ours	**0.806%**	**1.189%**	**1.009%**	**0.730%**	**1.130%**	**5.540%**	**2.112%**	**5.110%**	**1.860%**	**3.042%**

4.4. Performance Under Single-dataset Setting

We chose state-of-the-art methods for palmprint recognition to compare the performance of different palmprint datasets. Tab. 6 and Tab.7 present the EERs and accuracy for various methods across different datasets under a single-dataset setting.

The tables demonstrate that our method's EER and recognition accuracy for MS datasets approach the zero EER. Specifically:

- Our proposed method consistently achieved the best or near-best performance across different datasets and channels.
- In several cases, particularly for PolyU and various channels of MS, our method outperformed all others with an EER of 0% (accuracy of 100%).
- The method has exhibited robustness and effectiveness across diverse datasets and spectrums, signifying its versatility and superiority in palmprint recognition tasks.

To demonstrate the performance of various methods, we introduce ROC curves, presented in Fig.5. These curves compare the ROC metrics of different methods applied to the Tongji[10], PolyU [9], and IITD [11] datasets. As illustrated in Fig. 5, our method performs comparatively well.

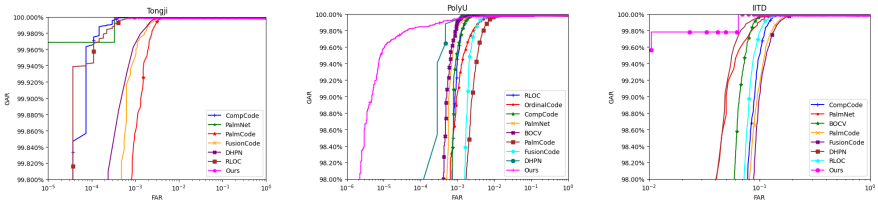

Figure 5: The ROC curves of different methods on different palmprint datasets. Among them, in Tongji's ROC image, ours is located at the top of the graph, almost overlapping the top of the vertical axis. This is because the FAR of our method is concentrated between 10^{-4} and 1, while there is almost no data between 0 and 10^{-4}.

Table 6: EERs of Single Dataset on Different Datasets(%).

Method \ Dataset	PolyU	Tongji	IITD
PalmCode[9]	0.3500	0.1100	5.4500
DHN[37]	0.0372	0.0879	4.3000
DHPN[38]	0.0302	0.0694	3.7316
PalmNet[25]	0.1110	0.0322	4.2000
W2ML[16]	-	1.7600	2.3300
C-LMCL^{Op}[40]	0.1250	0.2600	-
AML^{Op}[41]	0.8600	1.2100	1.7300
PKLNet[42]	-	0.8920	0.8200
CDR[43]	0.0010	-	-
V_{IR}^{Op}[44]	0.0800	-	0.4700
V_{MBFN}^{Op}[44]	0.1900	-	0.7400
CompNet[27]	0.0556	0.0250	0.5400
CO3Net[6]	0.0220	0.0050	0.4700
CCNet[5]	0.0004	0.0833	0.3547
Ours	**0**	**0.0015**	**0.2613**

Method \ Dataset	Red	Green	Blue	NIR
PalmCode[9]	0.2300	0.2500	0.2800	0.2000
DHN[37]	0.0380	0.0304	0.0403	0.0233
DHPN[38]	0.0369	0.0352	0.0213	0.0020
PalmNet[25]	0.0366	0.0087	0.0178	0.0871
DCPF^{Op}[39]	0.2672	0.1336	0.2672	0.1336
CompNet[27]	0	0.0055	0.0173	0.0025
CO3Net[6]	0	0.0009	0.0004	0.0005
CCNet[5]	0	0	0	0
Ours	**0**	**0**	**0**	**0**

Method \ Dataset	460	700	850
V_{IR}^{Op}[44]		0.5900 *	
V_{MBFN}^{Op}[44]		0.8500 *	
CO3Net[6]	0.6667	0.9444	0.4523
CCNet[5]	0.5625	0.6250	0.1667
Ours	**0.2552**	**0.5556**	**0.1667**

" * " indicates that the experimental result is obtained with the merge of 460, 700, and 850.

Table 7: Accuracies of Single Dataset on Different Datasets.

Method \ Dataset	PolyU	Tongji	IITD
W2ML[16]	-	93.39%	94.02%
C-LMCL^{Op}[40]	100%	99.93%	-
AML^{Op}[41]	99.02%	97.71%	99.02%
CDR[43]	100%	-	-
CompNet[27]	-	100%	98.77%
CO3Net[6]	100%	99.97%	98.26%
CCNet[5]	100%	99.87%	98.80%
Ours	**100%**	**100%**	**98.91%**

Method \ Dataset	Red	Green	Blue	NIR
CompNet[27]	99.98%	99.92%	99.94%	100%
CO3Net[6]	100%	99.98%	100%	100%
CCNet[5]	100%	100%	100%	100%
Ours	**100%**	**100%**	**100%**	**100%**

Method \ Dataset	460	700	850
CO3Net[6]	96.83%	96.83%	98.00%
CCNet[5]	98.50%	98.00%	98.75%
Ours	**99.33%**	**98.17%**	**99.17%**

4.5. Ablation Study

This section primarily examines the impact of the proposed method's components, such as histogram matching and Fourier alignment transform, and the role of various loss functions, including domain adversarial loss and feature similarity loss. To evaluate the contribution of each component, we selected the Tongji dataset [10] as the source domain D^0 and the PolyU dataset [9] as the target domain D^t. The results are presented in Tab. 8.

Results show that each component contributes significantly to performance improvement. Combining HM and FDA reduces the EER from 1.57% to 1.23% and increases accuracy from 99.76% to 99.79%. Adding SimLoss further improves accuracy to 99.87%, while incorporating DomLoss and SimLoss achieves the best performance with an EER of 1.13% and accuracy of 99.79%. This highlights the importance of all components in enhancing face recognition accuracy.

Table 8: Ablation Results with Different Components.

HM	FDA	L_{sim}	L_{adv}	*EER*	*Accuracy*
✗	✗	✗	✗	1.57%	99.76%
✔	✗	✗	✗	1.35%	99.87%
✗	✔	✗	✗	1.83%	99.29%
✔	✔	✗	✗	1.23%	99.79%
✔	✔	✔	✗	1.27%	99.87%
✔	✔	✔	✔	**1.13%**	**99.79%**

5. Conclusion

In this paper, we unveil the utiization of Fourier Alignment Transform and Histogram Matching for SSDG in palmprint recognition. Experimental results demonstrate a noticeable improvement in the model's recognition accuracy and generalization ability when combined with CCNet and incorporated with adversarial training. It

is noteworthy that the performance across different datasets, especially in cross-spectral testing across different datasets (such as PolyU, Tongji, and IITD), shows significant improvement. However, in the CasiaM dataset, our method is partially affected by the wavelength, leaving room for improvement in recognition accuracy in our future work.

Recognizing the resemblance between SSDG and zero-shot learning, our method can extend to open-set palmprint recognition, where the model must accurately classify palmprints belonging to previously unseen individuals. Since this method can be extended to target open-set palmprint recognition problems, it opens up new promising directions and goals for future research, which include not only high performance but also generalizability, simplicity, and efficiency.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Nos. 62376003) and Anhui Provincial Natural Science Foundation (No. 2308085MF200)

References

[1] L. Fei, G. Lu, W. Jia, S. Teng, and D. Zhang, "Feature extraction methods for palmprint recognition: A survey and evaluation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 2, pp. 346–363, 2018.

[2] K. Zhou, X. Zhou, L. Yu, L. Shen, and S. Yu, "Double biologically inspired transform network for robust palmprint recognition," *Neurocomputing*, vol. 337, pp. 24–45, 2019.

[3] D. Zhang, Z. Guo, G. Lu, L. Zhang, and W. Zuo, "An online system of multispectral palmprint verification," *IEEE transactions on instrumentation and measurement*, vol. 59, no. 2, pp. 480–490, 2009.

[4] R. Raghavendra and C. Busch, "Texture based features for robust palmprint recognition: a comparative study," *EURASIP Journal on Information Security*, vol. 2015, pp. 1–9, 2015.

[5] Z. Yang, H. Huangfu, L. Leng, B. Zhang, A. B. J. Teoh, and Y. Zhang, "Comprehensive competition mechanism in palmprint recognition," *IEEE Transactions on Information Forensics and Security*, 2023.

[6] Z. Yang, W. Xia, Y. Qiao, Z. Lu, B. Zhang, L. Leng, and Y. Zhang, "CO3Net: Coordinate-aware contrastive competitive neural network for palmprint recognition," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[7] L. Fei, W. K. Wong, S. Zhao, J. Wen, J. Zhu, and Y. Xu, "Learning spectrum-invariance representation for cross-spectral palmprint recognition," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023.

[8] K. Zhou, Z. Liu, Y. Qiao, T. Xiang, and C. C. Loy, "Domain generalization: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[9] D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 25, no. 9, pp. 1041–1050, 2003.

[10] L. Zhang, L. Li, A. Yang, Y. Shen, and M. Yang, "Towards contactless palmprint recognition: A novel device, a new benchmark, and a collaborative representation based identification approach," *Pattern Recognition*, vol. 69, pp. 199–212, 2017.

[11] A. Kumar and S. Shekhar, "Personal identification using multibiometrics rank-level fusion," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 5, pp. 743–752, 2010.

[12] L. Shen, Y. Zhang, K. Zhao, R. Zhang, and W. Shen, "Distribution alignment for cross-device palmprint recognition," *Pattern Recognition*, vol. 132, p. 108942, 2022.

[13] H. Shao, Y. Zou, C. Liu, Q. Guo, and D. Zhong, "Learning to generalize unseen

dataset for cross-dataset palmprint recognition," *IEEE Transactions on Information Forensics and Security*, 2024.

[14] Q. Xu, R. Zhang, Y.-Y. Wu, Y. Zhang, N. Liu, and Y. Wang, "Simde: A simple domain expansion approach for single-source domain generalization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 4797–4807.

[15] Y. Xian, C. H. Lampert, B. Schiele, and Z. Akata, "Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 9, pp. 2251–2265, 2018.

[16] H. Shao and D. Zhong, "Towards open-set touchless palmprint recognition via weight-based meta metric learning," *Pattern Recognition*, vol. 121, p. 108247, 2022.

[17] S. Li, R. Ma, J. Zhou, and B. Zhang, "Row-sparsity binary feature learning for open-set palmprint recognition," in *2022 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2022, pp. 1–8.

[18] Z. Su, K. Yao, X. Yang, K. Huang, Q. Wang, and J. Sun, "Rethinking data augmentation for single-source domain generalization in medical image segmentation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 2, 2023, pp. 2366–2374.

[19] Y. Zhang, W. Li, W. Sun, R. Tao, and Q. Du, "Single-source domain expansion network for cross-scene hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 32, pp. 1498–1512, 2023.

[20] Y. Yang and S. Soatto, "FDA: Fourier domain adaptation for semantic segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 4085–4095.

[21] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE transactions on pattern analysis and machine intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.

[22] A.-K. Kong and D. Zhang, "Competitive coding scheme for palmprint verification," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 1. IEEE, 2004, pp. 520–523.

[23] Z. Guo, D. Zhang, L. Zhang, and W. Zuo, "Palmprint verification using binary orientation co-occurrence vector," *Pattern Recognition Letters*, vol. 30, no. 13, pp. 1219–1227, 2009.

[24] Z. Yang, L. Leng, T. Wu, M. Li, and J. Chu, "Multi-order texture features for palmprint recognition," *Artificial Intelligence Review*, vol. 56, no. 2, pp. 995–1011, 2023.

[25] A. Genovese, V. Piuri, K. N. Plataniotis, and F. Scotti, "PalmNet: Gabor-PCA convolutional networks for touchless palmprint recognition," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3160–3174, Dec. 2019.

[26] S. Trabelsi, D. Samai, F. Dornaika, A. Benlamoudi, K. Bensid, and A. Taleb-Ahmed, "Efficient palmprint biometric identification systems using deep learning and feature selection methods," *Neural Computing and Applications*, vol. 34, no. 14, pp. 12 119–12 141, 2022.

[27] X. Liang, J. Yang, G. Lu, and D. Zhang, "Compnet: Competitive neural network for palmprint recognition using learnable gabor kernels," *IEEE Signal Processing Letters*, vol. 28, pp. 1739–1743, 2021.

[28] X. Du, D. Zhong, and H. Shao, "Cross-domain palmprint recognition via regularized adversarial domain adaptive hashing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 6, pp. 2372–2385, 2020.

[29] H. Shao and D. Zhong, "Towards cross-dataset palmprint recognition via joint pixel and feature alignment," *IEEE Transactions on Image Processing*, vol. 30, pp. 3764–3777, 2021.

[30] ——, "Multi-target cross-dataset palmprint recognition via distilling from multi-teachers," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[31] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4690–4699.

[32] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in neural information processing systems*, vol. 33, pp. 18 661–18 673, 2020.

[33] J. Wang, C. Lan, C. Liu, Y. Ouyang, T. Qin, W. Lu, Y. Chen, W. Zeng, and S. Y. Philip, "Generalizing to unseen domains: A survey on domain generalization," *IEEE transactions on knowledge and data engineering*, vol. 35, no. 8, pp. 8052–8072, 2022.

[34] Chinese Academy of Sciences, Inst. Automation. (2005) Biometrics Ideal Test. [Accessed: 28-Jan-2016]. [Online]. Available: http://biometrics.idealtest.org/dbDetailForUser.do?id=5

[35] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7167–7176.

[36] W. Lu, J. Wang, H. Li, Y. Chen, and X. Xie, "Domain-invariant feature exploration for domain generalization," *arXiv preprint arXiv:2207.12020*, 2022.

[37] T. Wu, L. Leng, M. K. Khan, and F. A. Khan, "Palmprint-palmvein fusion

recognition based on deep hashing network," *IEEE Access*, vol. 9, pp. 135 816–135 827, 2021.

[38] D. Zhong, S. Liu, W. Wang, and X. Du, "Palm vein recognition with deep hashing network," in *Proc. Chin. Conf. Pattern Recognit. Comput. Vis. (PRCV)*, 2018, pp. 38–49.

[39] S. Trabelsi, D. Samai, A. Meraoumia, K. Bensid, and A. Taleb-Ahmed, "An improved multispectral palmprint system using deep cnn-based palm-features," in *2019 International Conference on Advanced Electrical Engineering (ICAEE)*. IEEE, 2019, pp. 1–6.

[40] D. Zhong and J. Zhu, "Centralized large margin cosine loss for open-set deep palmprint recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 6, pp. 1559–1568, 2019.

[41] J. Zhu, D. Zhong, and K. Luo, "Boosting unconstrained palmprint recognition with adversarial metric learning," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 2, no. 4, pp. 388–398, 2020.

[42] X. Liang, D. Fan, J. Yang, W. Jia, G. Lu, and D. Zhang, "PKLNet: Keypoint localization neural network for touchless palmprint recognition based on edge-aware regression," *IEEE Journal of Selected Topics in Signal Processing*, 2023.

[43] W. Jia, B. Zhang, J. Lu, Y. Zhu, Y. Zhao, W. Zuo, and H. Ling, "Palmprint recognition based on complete direction representation," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4483–4498, 2017.

[44] Y. Zhang, L. Zhang, R. Zhang, S. Li, J. Li, and F. Huang, "Towards palmprint verification on smartphones," *arXiv preprint arXiv:2003.13266*, 2020.