

# AdaptPFL: Unlocking Cross-Device Palmprint Recognition via Adaptive Personalized Federated Learning with Feature Decoupling

Zirui Zhang<sup>1</sup>, Donghai Guan<sup>1,\*</sup>, Çetin Kaya Koç<sup>1</sup>, Jie Wen<sup>2</sup>, Qi Zhu<sup>1,\*</sup>

<sup>1</sup>Nanjing University of Aeronautics and Astronautics

<sup>2</sup>Harbin Institute of Technology, Shenzhen

zhangzirui@nuaa.edu.cn, dhguan@nuaa.edu.cn, cetinkoc@ucsb.edu, wenjie@hit.edu.cn, zhuqi@nuaa.edu.cn

## Abstract

Contactless palmprint recognition has recently emerged as a promising biometric technology. However, traditional methods that require sharing user data introduce substantial security risks. While federated learning offers privacy-preserving solutions, it often compromises recognition accuracy due to feature distribution drift caused by external factors such as lighting and devices. To address this issue, we propose an **adaptive personalized federated learning** framework (AdaptPFL). The central innovation lies in decomposing palmprint features into identity-related and contextual-related components using a feature decoupling mechanism. This design isolates the influence of external environmental factors on identity recognition through de-entanglement. Furthermore, two adaptive aggregation strategies are introduced to correct client drift: (1) **Intra-Local Adaptive Aggregation (ILAA)**, which addresses intra-client drift by adaptively combining the two decoupled feature types; (2) **Global-Local Adaptive Aggregation (GLAA)**, which corrects inter-client drift by adaptively aggregating model parameters. Experimental results demonstrate that AdaptPFL achieves superior performance compared to existing state-of-the-art methods.

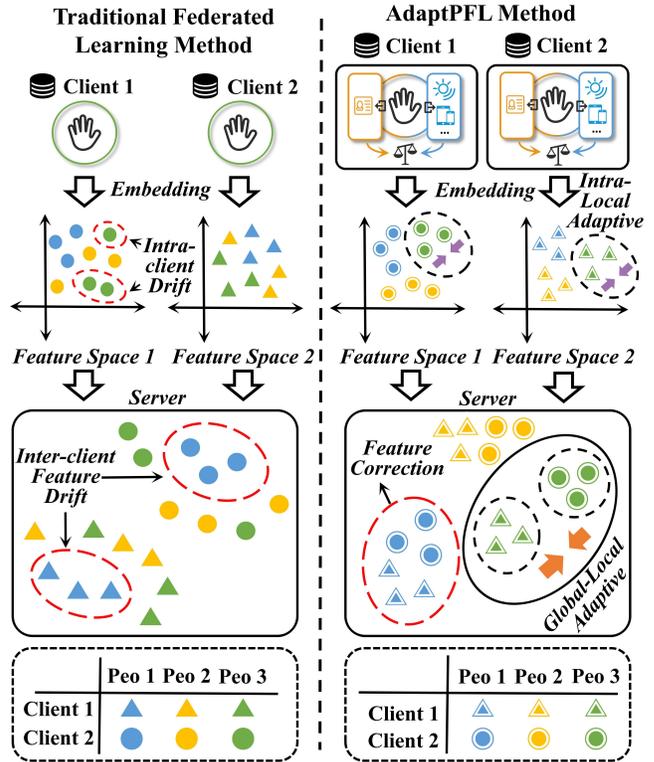


Figure 1: Example of client-side drift

## 1 Introduction

Over the past two decades, biometric identification technologies have garnered significant attention and have been widely applied across various domains. Among these, face recognition stands out for its convenience and widespread use. However, the ability to capture facial images without the user’s consent has raised growing concerns regarding potential privacy breaches. In contrast, palmprint recognition typically requires active user cooperation to capture palmprint images. As a result, palmprint recognition is considered a more privacy-preserving biometric technology than face recognition.

Although palmprint recognition offers better privacy protection than face recognition, traditional methods still pose significant privacy and security risks in many mobile application scenarios requiring personal data sharing. As a result, researchers have been increasing interest in leveraging federated learning [Li *et al.*, 2020a] to safeguard user privacy and security. The fundamental principle of federated learning is to decentralize the training of machine learning models by shifting the computation from a central server to various devices or data centres. Each data holder trains the model locally and shares only the model parameters or updates rather than raw data. This approach enables the participation of sensitive data in the training process without leaving the local domain, thus ensuring user data security and privacy protection.

\* Corresponding authors

Despite traditional federated learning methods offering a degree of privacy and security for users, complex application scenarios in palmprint recognition continue to pose significant challenges. As illustrated on the left side of Fig. 1, the same people in the traditional method produce obvious feature drift between different clients. There is still some drift even within the same client. These discrepancies lead to aggregation challenges at the server, ultimately reducing recognition accuracy. To address these issues, we propose an adaptive personalized federated learning framework based on feature decoupling for cross-device palmprint recognition. This method decouples palmprint features into identity and contextual information, as shown on the right side of Fig. 1. An intra-local adaptive aggregation strategy is introduced to correct intra-client drift, effectively minimizing the impact of external environmental factors. Furthermore, an adaptive global-local aggregation strategy addresses inter-client drift by integrating global information essential for client models.

The contribution of this paper can be summarised as follows:

(i) We design a de-entanglement mechanism using a deep edge detection model and a mutual information method. The influence of external factors on identification is isolated by decoupling the palmprint information.

(ii) Two adaptive aggregation strategies, ILAA and GLAA, are introduced to correct intra-client drift and inter-client drift, respectively.

(iii) We conducted extensive experiments on three contactless palmprint databases. The results demonstrate the method's higher accuracy than the SOTA methods and effectively mitigate the client-side drift problem.

## 2 Related Work

The section focuses on the current work on palmprint recognition and federated learning in detail.

### 2.1 Palmprint Recognition

In recent years, deep learning-based methods have been applied to palmprint recognition. Genovese et al. [Genovese et al., 2019] proposed PalmNet for unsupervised palmprint recognition. It includes a convolutional neural network, Gabor filter, and principal component analysis. Zhong et al. [Zhong and Zhu, 2019] proposed an end-to-end palmprint recognition method, which applies a new loss function of CNN to enhance the inner class compactness of palmprint features. Matkowski et al. [Matkowski et al., 2019] proposed the palmprint recognition network EE-PRnet, which consists of two main networks that can solve the palmprint recognition problem in uncontrollable and non-collaborative environments. Zhao et al. [Zhao et al., 2019] proposed a joint deep convolutional feature representation applied to feature extraction of hyperspectral palmprint images. Jia et al. [Jia et al., 2022] improved MobileNet-V3 and proposed a lightweight convolutional neural network for palmprint recognition. Zhao et al. [Zhao and Zhang, 2020] used deep convolutional neural networks to learn complete and discriminative convolutional features and proposed a jointly constrained least squares regression framework. This framework combined deep local convolutional features to solve the

undersampling classification problem in palmprint recognition. Shao et al. [Shao and Zhong, 2021] introduced a deep hash network of MobileFaceNets to extract discriminative features. It improves the matching efficiency and allows the target network to adapt to unlabelled target palmprint images. Fan et al. [Fan et al., 2022] proposed using deep learning architecture to extract the main lines of palmprints to reduce the effect of fine lines and further classify the palmprint phenotypes from the 2D palmprint images. Samai et al. [Samai et al., 2018] proposed using DCTNet for 3D palmprint recognition. Chaa et al. [Chaa et al., 2019] first used a single-scale retina algorithm to enhance the depth image of 3D palmprints and then used PCANet for recognition.

However, most contactless palmprint recognition methods ignore user privacy issues. The huge security risk of sharing personal data among different users is also an urgent problem to solve.

### 2.2 Federated Learning

Currently, the federated learning technique is a distributed machine learning framework that effectively solves the data silo problem. McMahan et al. [McMahan et al., 2017] proposed a FedAvg algorithm, an important milestone in developing federated learning. Most of the current federated learning methods are based on the idea of the FedAvg algorithm and continuously optimize and upgrade the federated learning algorithm. Li et al. [Li et al., 2020b] proposed the FedProx algorithm, which adds a proximal term to the client's original loss function. This maintains the similarity between each client's local and global model parameters in federated learning. The robustness of the model is improved by balancing the differences between the global and local models. With the continuous development of federated learning, techniques such as comparative learning, prototype learning, and meta-learning have been introduced into the research and practice of federated learning. Li et al. [Li et al., 2021] proposed the MOON algorithm, which performs comparative learning at the model level by comparing representations learned from different models. The aim is to reduce the distance between the local and central weight representations. Thereby increasing the distance to the local weight representation of the previous round. The similarity of model representations is used to correct the bias in local training and alleviate the data heterogeneity problem. Mu et al. [Mu et al., 2023] introduced prototype learning based on the MOON algorithm, which uses prototypes as global knowledge to correct the bias in local training for each client. Decrease the distance between the prototype of a local class and the global prototype of the same class. Thus increasing the distance from the global prototype of a different class.

The above approaches aim to train a global model with generalization performance for all clients. However, when the data distribution of cross-device clients varies significantly by conditions such as lighting and brand, a single global model may not be able to meet the diverse needs of each client [Huang et al., 2021]. Therefore, this paper presents two adaptive aggregation strategies for intra-client and inter-client drift phenomena.

### 3 Methodology

The section describes each module in the model in detail. It mainly contains a feature decoupling module, a classifier alignment module, an intra-local adaptive module, and a global-local adaptive module. Fig. 2 shows the architecture of the federated model AdaptPFL.

#### 3.1 Feature Decoupling Module

We decouple palmprint features into two non-overlapping and independent parts: identity and contextual information. The identity information feature contains the unique feature identifier of the personal information in the data, while contextual information features contain other features besides personal information. The next part of the section describes in detail how these two types of features are extracted.

##### Identity-Related Extraction

This paper focuses on extracting identity-related features from palmprint images. The key identity information is primarily contained within the three main palm lines. We employ edge detection methods to extract these features that effectively capture identity information from the primary palm lines. Additionally, the resulting grayscale image helps mitigate the influence of lighting conditions, further enhancing the robustness of the feature extraction process.

We use a deep edge detection algorithm to accomplish feature extraction of identity information. A deep convolutional neural network consists of many layers, and different layers can learn various levels of features. Shallow layers can extract basic physical features, and advanced features, such as trunk lines, can be extracted as the layer depth increases. Since training an edge detection model requires a large amount of labelled data, we first train the model on a publicly available benchmark edge detection dataset. The best learning parameters are saved after the training is completed, and then the parameters are introduced for training using the migration learning method. The main palm line feature is extracted by fine-tuning it, and this feature is used to contain identity information.

Inspired by the HED [Xie and Tu, 2015] method, the architecture of the edge detection algorithm chosen is shown in the identity-related extraction module on the top right of Fig. 2. The model uses the VGG16 network as the backbone network, and the specific training process contains two phases. The first phase receives information through five different-sized receiving fields. The second stage connects a side-output layer  $i$  to the last convolutional layer of each receptive field to learn multi-scale features. Finally, the outputs of the five side-output layers are connected as inputs to the final output layer to achieve edge detection.

For simplicity, we set of all network layer parameters is denoted as  $W$ . The network has five side output layers, each of which is also associated with a classifier. The corresponding weights are denoted as  $w = (w^{(1)}, \dots, w^{(5)})$ , and the corresponding fusion weights are denoted as  $h = (h_1, \dots, h_5)$ . Its specific computational procedure is shown below:

$$\left( \hat{H}_{side}^{(1)}, \hat{H}_{side}^{(2)}, \dots, \hat{H}_{side}^{(5)} \right) = \text{VGG16}(X, (W, w, h)) \quad (1)$$

$$X_i = \text{Average} \left( \hat{H}_{side}^{(1)}, \hat{H}_{side}^{(2)}, \dots, \hat{H}_{side}^{(5)} \right) \quad (2)$$

#### Contextual-Related Extraction

This paper introduces a division strategy to extract the features of contextual information. The strategy divides the data feature  $X$  into two parts in the same metric space, making  $X = X_i + X_c$ . Specifically, the data feature  $X \in \mathbb{R}^d$  with dimension  $d$  is decomposed into the identity information feature  $X_i \in \mathbb{R}^d$  and the contextual information feature  $X_c \in \mathbb{R}^d$ . Intuitively, the identity information feature contains all the labelled information. In contrast, the contextual information feature contains all the data except the labelled information.

As shown in the contextual-related extraction module in the lower right of Fig. 2, the edge detection method has obtained the initial features of personal identity information. We use mutual information to generate contextual features  $X_c$  for the data  $X$ . This paper aims to ensure that the features after removing the personal identification information are close to the original features. More specifically, the contextual information features are obtained with the goal of:

$$\min_{X_i} I(X; Y | X_i), \text{ s.t. } I(X; X - X_i | X_i) \geq I_{FF} \quad (3)$$

where  $Y$  is the label of  $X$ , and  $X_i$  should contain necessary information about the label  $Y$ .  $X - X_i$  denotes the contextual information feature  $X_c$ , and  $I_{FF}$  is a constant. With this strategy, we can obtain both contextual information features  $X_c$  and optimized identity information features  $X_i$  at the same time.

#### 3.2 Classifier Alignment Module

We propose a classifier module that aligns identity and contextual information. It also designs a maximum mean difference (mmd) loss function.

##### Classifier Structure Design

The classifier alignment module is a multi-output network consisting of two specific predictors for predicting the outcome of identity information feature  $X_i$  and contextual information feature  $X_c$ , respectively. A ResNet18 network forms each predictor. Each predictor is a softmax classifier that makes a classification judgment after the two feature extractors extract the features to obtain  $P_i$  and  $P_c$ , respectively. Its specific computation is shown below:

$$P_i = \text{softmax}(\text{ResNet18}(X_i)) \quad (4)$$

$$P_c = \text{softmax}(\text{ResNet18}(X_c)) \quad (5)$$

##### mmd Loss Function Design

Since the classifiers are trained independently, their predictions of target samples may diverge, especially for target samples near class boundaries. Intuitively, different classifiers should get the same prediction when receiving different features from the same sample. Therefore, this paper aims to minimize the differences between all classifiers.

To align the distribution of predicted outcomes for identity and contextual information features, we choose mmd as an estimate of the difference in the distribution of predicted outcomes for the two features. mmd is based on the idea that all statistical data are the same if the generated distributions are

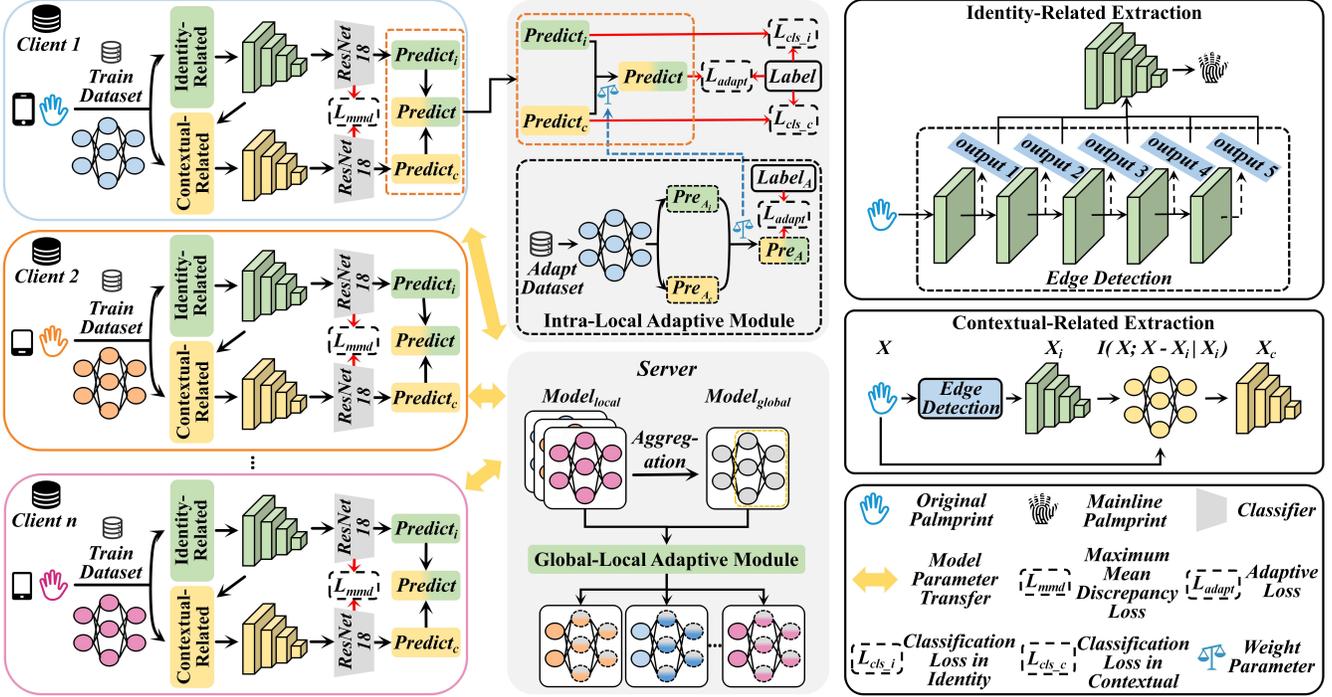


Figure 2: The overall architecture of the proposed method. It mainly contains a feature decoupling module, a classifier alignment module, an intra-local adaptive module, and a global-local adaptive module.

the same. Formally, mmd defines the following measure of variation:

$$D_H(p, q) = \|E_p[\phi(X_i)] - E_q[\phi(X_c)]\|_H^2 \quad (6)$$

where  $H$  is the reproducing kernel Hilbert space (RKHS) with feature kernel  $k$ , where  $\phi(\cdot)$  denotes some feature mapping that maps the original sample to the RKHS.  $k$  denotes  $k(X_i, X_c) = \langle \phi(X_i), \phi(X_c) \rangle$ , where  $\langle \cdot, \cdot \rangle$  denotes the inner product of the vectors. We use Eq. (6) to estimate the difference between the predictions of the two classifiers. The formula for the  $L_{mmd}$  loss is shown below:

$$L_{mmd} = D(P_i, P_c) \quad (7)$$

Each classifier can reduce the variance between classifier predictions by minimizing Eq. (7).

### 3.3 Intra-Local Adaptive Module

Consider that  $X_i$  contains important identity information and  $X_c$  contains rich auxiliary information. Therefore, we use the weighted average of the predictions  $P_i$  and  $P_c$  generated by  $X_i$  and  $X_c$  for inference. We introduce a personalization weight  $\lambda_i \in [0, 1]$  for the  $i$ -th client, where the weight difference increases personalization. The client uses  $P$  for integrated prediction, which is computed as shown below:

$$P = \lambda P_i + (1 - \lambda) P_c \quad (8)$$

However, the optimal value of  $\lambda$  may differ for each client due to client drift issues. Therefore, we propose to learn an appropriate value of  $\lambda_i$  for each client instead of setting a common value for it centrally. We randomly split a small

subset from the training set  $D_T$  to train the  $i$ -th client for  $\lambda_i$ , name it the adaptation set and denote it by  $D_A^\lambda$ . We learn  $\lambda_i$  by training on the adaptive set, which is computed using the cross-entropy loss function. Given the predictions generated by the identity information model and the contextual information model, the learning process of the loss function  $L_{adapt}$  and the learning rate  $\eta$  on the adaptive set are denoted as:

$$L_{adapt} = L_{CE}(P, y) \quad (9)$$

$$\lambda_i = \lambda_i - \eta \frac{\partial L_{adapt}}{\partial \lambda_i} \quad (10)$$

The algorithm uses stochastic gradient descent (SGD) for small batch iterations while sandwiching  $\lambda$  in  $[0, 1]$ .

### 3.4 Global-Local Adaptive Module

While traditional aggregation methods can generate a global model, the global model performs poorly in generalization on a per-client basis. Inspired by FedALA [Zhang *et al.*, 2023], this paper proposes an adaptive aggregation method between global-local models. The method judiciously aggregates global and local models to fit local goals. We show the adaptive aggregation method's learning process in Fig. 3.

In the  $r$ -th iteration of traditional federated learning, the server sends the old global model  $\theta^{r-1}$  to the client  $k$ .  $\theta^{r-1}$  overwrites the old local model  $\theta_k^{r-1}$  to get the initialized local model  $\hat{\theta}_k^r$  for local model training, i.e.,  $\hat{\theta}_k^r = \theta^{r-1}$ . In contrast, we propose an adaptive method to aggregate the global and local models instead of overwriting. The specific calcu-

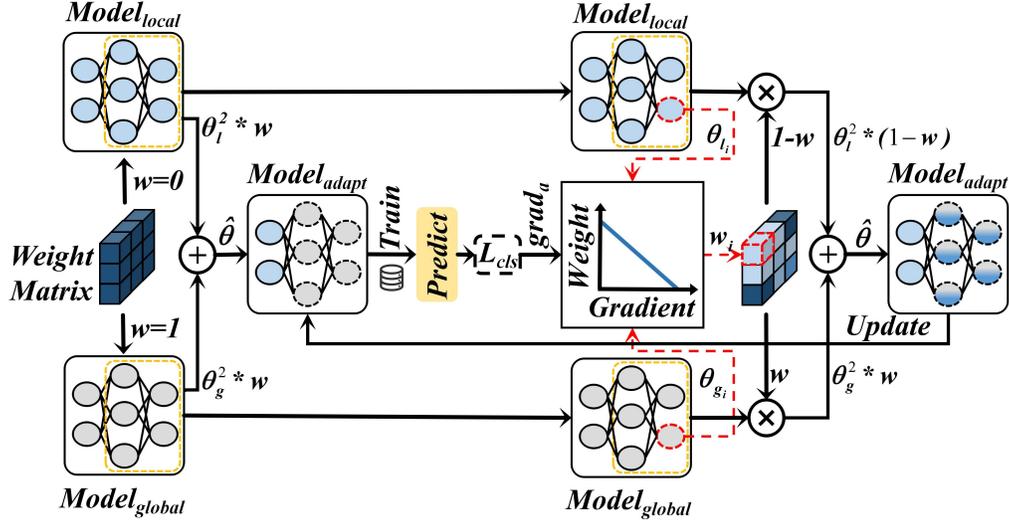


Figure 3: Global-Local Adaptive Module architecture diagram in our proposed method.

lation method is shown as follows:

$$\hat{\theta}_k^r = \theta_k^{r-1} \odot W_{k,1} + \theta^{r-1} \odot W_{k,2} \quad (11)$$

where  $\odot$  is the Hadamard product,  $W_{k,1}$  and  $W_{k,2}$  are the learnable weight matrices and  $W_{k,1} + W_{k,2} = 1$ . For coverage,  $W_{k,1} = 0$  and  $W_{k,2} = 1$ .

However, learning  $W_{k,1}$  and  $W_{k,2}$  with constraints by the gradient-based learning method is difficult. Therefore, we merge  $W_{k,1}$  and  $W_{k,2}$ , and the modification to Eq. (11) is shown below:

$$\hat{\theta}_k^r = \theta_k^{r-1} + (\theta^{r-1} - \theta_k^{r-1}) \odot W_k \quad (12)$$

where  $W_k$  is the learnable weight matrix.

### 3.5 Overall Loss Function Design

The loss function  $L_{total}$  of the method consists of three parts: classification loss  $L_{cls,i}$  and  $L_{cls,c}$  for the identity information module and contextual information module,  $L_{mmd}$  loss, and  $L_{adapt}$  loss. We use the cross-entropy loss as the classification loss for the two classifiers, as shown in Eq. (13) and Eq. (14). By minimizing the classification loss, the network can classify accurately. By minimizing the mmd loss, the network can reduce the differences between the classifiers. The network can better integrate the identity and contextual information modules by minimizing the  $L_{adapt}$  loss. The formula for the total loss  $L_{total}$  is shown below:

$$L_{cls,i} = L_{CE}(P_i, y_i) \quad (13)$$

$$L_{cls,c} = L_{CE}(P_c, y_i) \quad (14)$$

$$L_{total} = L_{cls,i} + L_{cls,c} + \beta L_{mmd} + L_{adapt} \quad (15)$$

where  $\beta$  is a hyperparameter, training is mainly done using the standard SGD algorithm.

## 4 Experiment

We evaluate the method's effectiveness on three public non-contact palmprint datasets. The method is compared with SOTA models, and the basic components of the models are further analyzed.

### 4.1 Datasets

To evaluate the method's recognition accuracy, extensive experiments and analyses were conducted on three contactless palmprint databases, XJTU-UP [Shao *et al.*, 2019], MPD [Zhang *et al.*, 2020], and Multi-Spectral [Zhang *et al.*, 2009]. Table 1 summarises the details of the three datasets.

**XJTU-UP:** An unconstrained palmprint database collected from two smartphones, iPhone 6S and Huawei Mate8. A total of four datasets are included, which are noted as HuaWei Nature (HWN), HuaWei Flash (HWF), iPhone Nature (IPN) and iPhone Flash (IPF).

**MPD:** Palm images are collected by two smartphones, HuaWei and XiaoMi, denoted as HW and XM, respectively.

**Multi-Spectral:** Palmprint images are taken from different spectral bands, including blue, green, red, and near-infrared.

### 4.2 Comparison of Other Methods

We validate the effectiveness of the methods through several experiments. Three baseline methods are included: FedAvg [McMahan *et al.*, 2017], FedProx [Li *et al.*, 2020b], and SCAFFOLD [Karimireddy *et al.*, 2020]. Three SOTA methods are included: FedAPEN [Qin *et al.*, 2023], FedFed [Yang *et al.*, 2024], and FedML [Shao *et al.*, 2024].

As can be seen from the experimental results in Table 2, our method produces excellent performance on all three datasets. In particular, on the XJTU-UP dataset, which is poorly recognized by the SOTA method, our method's F1 value can still achieve excellent performance of 97.17%. Some SOTA methods have also produced better performance on the multi-spectral dataset. This is because Multi-Spectral images capture different bands of light, some of which (e.g., near-infrared light) are less sensitive to changes in ambient light. Therefore, changes in the external environment have less effect on them and produce better performance. However, the AdaptPFL method still shows the best performance on this dataset.

Dataset	XJTU-UP				MPD		Multi-Spectral			
	HWN	HWF	IPN	IPF	HW	XM	Blue	Green	Red	NIR
Number of volunteers	100	100	100	100	200	200	500	500	500	500
Image number per palm	10	10	10	10	20	20	12	12	12	12
Total number of palms	2000	2000	2000	2000	8000	8000	6000	6000	6000	6000
Example										

Table 1: Details of the three palmprint datasets

Dataset	Method	Accuracy(%)	Precision(%)	Recall(%)	F1(%)	Specificity(%)	FPR(%)
XJTU-UP	FedAvg	77.98	82.86	77.65	78.58	99.89	0.11
	FedProx	67.58	78.37	67.30	69.74	99.84	0.16
	SCAFFOLD	81.22	84.35	81.27	81.11	99.90	0.10
	FedAPEN	88.56	87.92	84.36	86.14	99.93	0.07
	FedFed	93.43	93.57	92.36	92.97	99.98	0.02
	FedML	88.75	89.21	87.36	88.29	99.96	0.04
	<b>AdaptPFL</b>	<b>97.21</b>	<b>97.74</b>	<b>97.22</b>	<b>97.17</b>	<b>99.99</b>	<b>0.01</b>
MPD	FedAvg	92.19	93.32	92.19	92.25	99.96	0.04
	FedProx	88.50	90.26	88.50	88.33	99.94	0.06
	SCAFFOLD	74.75	78.53	74.74	74.11	99.87	0.13
	FedAPEN	94.32	93.68	95.36	94.52	99.97	0.03
	FedFed	96.84	96.25	97.16	96.71	99.98	0.02
	FedML	95.34	94.98	93.17	94.08	99.97	0.03
	<b>AdaptPFL</b>	<b>98.75</b>	<b>99.03</b>	<b>98.75</b>	<b>98.74</b>	<b>99.99</b>	<b>0.01</b>
Multi-Spectral	FedAvg	89.00	93.75	89.15	90.00	99.98	0.02
	FedProx	88.67	93.88	88.67	89.86	99.98	0.02
	SCAFFOLD	84.75	88.76	84.78	85.29	99.97	0.03
	FedAPEN	97.62	97.66	98.19	97.93	99.99	0.01
	FedFed	99.11	98.98	98.36	98.67	99.99	0.01
	FedML	97.56	97.86	98.02	97.94	99.99	0.01
	<b>AdaptPFL</b>	<b>99.79</b>	<b>99.80</b>	<b>99.83</b>	<b>99.79</b>	<b>100.00</b>	<b>0.00</b>

Table 2: Comparison results of baseline and SOTA methods

### 4.3 Ablation Study

We conduct ablation studies on three contactless palmprint datasets to validate the effectiveness of the method’s core modules. The results are shown in Table 3.

method	XJTU-UP	MPD	Multi-Spectral
AdaptPFL	97.21	98.75	99.79
-IIM	93.72	94.26	96.24
-CIM	96.22	97.11	98.51
-ILAA	96.16	97.43	98.36
-GLAA	95.47	96.34	97.69

Table 3: Results of ablation studies on three datasets

This paper conducts four types of ablation studies: removal of the Identity Information Module (-IIM), removal of the Contextual Information Module (-CIM), removal of the Intra-Local Adaptive Aggregation (-ILAA), and removal of

the Global-Local Adaptive Aggregation (-GLAA). The experimental results show that the method’s performance is degraded to different degrees when the core modules are removed. In particular, the method’s performance degrades most severely when we remove the identity information module after detangling. This indicates that the decoupling mechanism we designed effectively separates the identity information from the palmprint features. The performance degradation is more severe after GLAA removal compared to ILAA. This is because the inter-client drift phenomenon is much larger than the intra-client drift phenomenon. The results of the ablation studies show that all our core modules are effective.

### 4.4 Histogram Analysis of Feature Distribution in Different Conditions

The section verifies the method’s validity by randomly selecting the distribution of palmprint features generated by a batch of volunteers under different scenes. Specifically, by

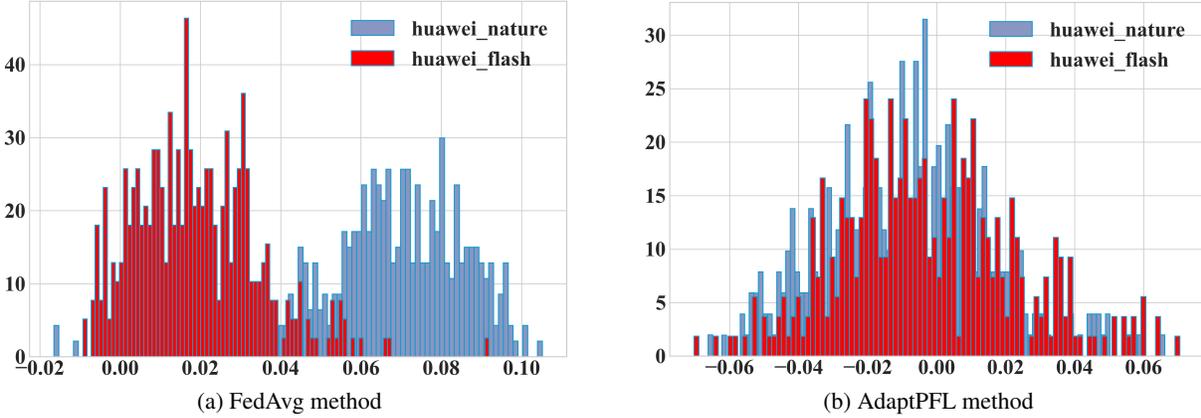


Figure 4: Histograms of different lighting in the dataset XJTU-UP

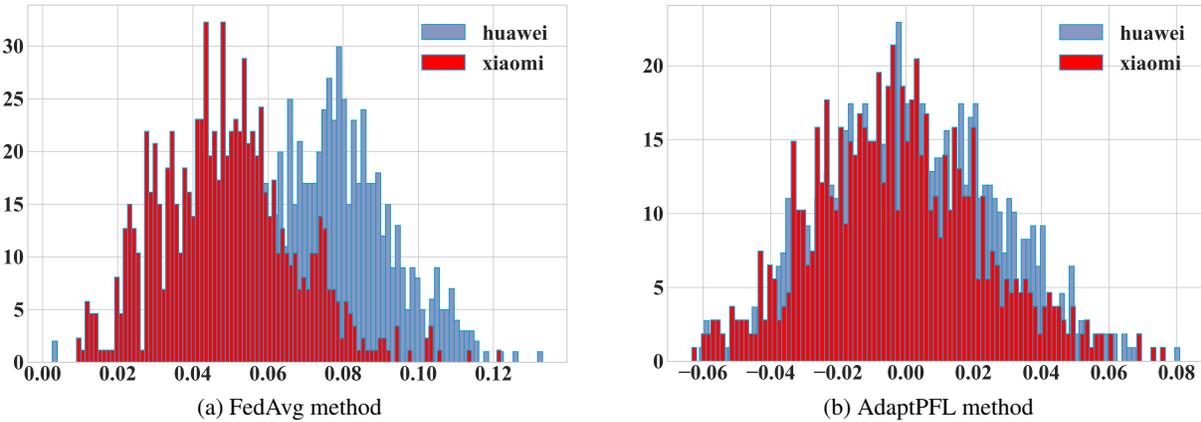


Figure 5: Histograms of different brands in the dataset MPD

summing and averaging the palmprint features of each volunteer in various scenarios as the current valid features, the palmprint feature distribution of the same batch of volunteers in different scenarios is formed. We mainly verify the effect of the drift of two scenes, mobile phone brand and light, on the feature distribution. The specific experimental results are presented in Figs. 4 and 5.

The experimental results show that the traditional federated learning method using different mobile phone brands or light conditions produced two different palmprint feature distributions for the same group of volunteers. This indicates that traditional federated learning is susceptible to the influence of external conditions to produce client drift. The present method produces similar feature distributions in different scenarios, which indicates that it can effectively mitigate the client-side drift phenomenon produced by external factors.

## 5 Conclusion

This paper proposes an adaptive personalized federated learning method based on feature decoupling for cross-device

palmprint recognition. The method designs a decoupling mechanism to decompose palmprint features into identity and contextual information, effectively isolating the mutual influence between external factors and identification. Meanwhile, we present two adaptive aggregation strategies for intra-client and inter-client drift phenomena, respectively. They used an intra-local adaptive module and a global-local adaptive module to correct the drift phenomenon further. We verify the method’s effectiveness on three public non-contact palmprint datasets. The results show that the present model performs better than SOTA methods.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (Nos. 62472220, 62371234, 62076129), Jiangsu Province 100 Foreign Experts Introduction Plan (BX2022012), Natural Science Foundation of Jiangsu Province (No. BK20231438), and Key Research and Development Plan of Jiangsu Province (No. BE2022842).

## References

- [Chaa *et al.*, 2019] Mourad Chaa, Zahid Akhtar, and Abdelouahab Attia. 3d palmprint recognition using unsupervised convolutional deep learning network and svm classifier. *IET Image Processing*, 13(5):736–745, 2019.
- [Fan *et al.*, 2022] Yu Fan, Jinxi Li, Shaoying Song, Haiguo Zhang, Sijia Wang, and Guangtao Zhai. Palmprint phenotype feature extraction and classification based on deep learning. *Phenomics*, 2(4):219–229, 2022.
- [Genovese *et al.*, 2019] Angelo Genovese, Vincenzo Piuri, Konstantinos N Plataniotis, and Fabio Scotti. Palmnet: Gabor-pca convolutional networks for touchless palmprint recognition. *IEEE Transactions on Information Forensics and Security*, 14(12):3160–3174, 2019.
- [Huang *et al.*, 2021] Yutao Huang, Lingyang Chu, Zirui Zhou, Lanjun Wang, Jiangchuan Liu, Jian Pei, and Yong Zhang. Personalized cross-silo federated learning on non-iid data. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 7865–7873, 2021.
- [Jia *et al.*, 2022] Wei Jia, Qiang Ren, Yang Zhao, Shujie Li, Hai Min, and Yanxiang Chen. Eepnet: An efficient and effective convolutional neural network for palmprint recognition. *Pattern Recognition Letters*, 159:140–149, 2022.
- [Karimireddy *et al.*, 2020] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for federated learning. In *International conference on machine learning*, pages 5132–5143. PMLR, 2020.
- [Li *et al.*, 2020a] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60, 2020.
- [Li *et al.*, 2020b] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450, 2020.
- [Li *et al.*, 2021] Qinbin Li, Bingsheng He, and Dawn Song. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10713–10722, 2021.
- [Matkowski *et al.*, 2019] Wojciech Michal Matkowski, Tingting Chai, and Adams Wai Kin Kong. Palmprint recognition in uncontrolled and uncooperative environment. *IEEE Transactions on Information Forensics and Security*, 15:1601–1615, 2019.
- [McMahan *et al.*, 2017] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agueria y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.
- [Mu *et al.*, 2023] Xutong Mu, Yulong Shen, Ke Cheng, Xueli Geng, Jiakuan Fu, Tao Zhang, and Zhiwei Zhang. Fedproc: Prototypical contrastive federated learning on non-iid data. *Future Generation Computer Systems*, 143:93–104, 2023.
- [Qin *et al.*, 2023] Zhen Qin, Shuiguang Deng, Mingyu Zhao, and Xueqiang Yan. Fedapen: personalized cross-silo federated learning with adaptability to statistical heterogeneity. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1954–1964, 2023.
- [Samai *et al.*, 2018] Djamel Samai, Khaled Bensid, Abdallah Meraoumia, Abdelmalik Taleb-Ahmed, and Mouldi Bedda. 2d and 3d palmprint recognition using deep learning method. In *2018 3rd international conference on pattern analysis and intelligent systems (PAIS)*, pages 1–6. IEEE, 2018.
- [Shao and Zhong, 2021] Huikai Shao and Dexing Zhong. One-shot cross-dataset palmprint recognition via adversarial domain adaptation. *Neurocomputing*, 432:288–299, 2021.
- [Shao *et al.*, 2019] Huikai Shao, Dexing Zhong, and Xuefeng Du. Efficient deep palmprint recognition via distilled hashing coding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 0–0, 2019.
- [Shao *et al.*, 2024] Huikai Shao, Chengcheng Liu, Xiaojiang Li, and Dexing Zhong. Privacy preserving palmprint recognition via federated metric learning. *IEEE Transactions on Information Forensics and Security*, 2024.
- [Xie and Tu, 2015] Saining Xie and Zhuowen Tu. Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision*, pages 1395–1403, 2015.
- [Yang *et al.*, 2024] Zhiqin Yang, Yonggang Zhang, Yu Zheng, Xinmei Tian, Hao Peng, Tongliang Liu, and Bo Han. Fedfed: Feature distillation against data heterogeneity in federated learning. *Advances in Neural Information Processing Systems*, 36, 2024.
- [Zhang *et al.*, 2009] David Zhang, Zhenhua Guo, Guangming Lu, Lei Zhang, and Wangmeng Zuo. An online system of multispectral palmprint verification. *IEEE transactions on instrumentation and measurement*, 59(2):480–490, 2009.
- [Zhang *et al.*, 2020] Yingyi Zhang, Lin Zhang, Ruixin Zhang, Shaoxin Li, Jilin Li, and Feiyue Huang. Towards palmprint verification on smartphones. *arXiv preprint arXiv:2003.13266*, 2020.
- [Zhang *et al.*, 2023] Jianqing Zhang, Yang Hua, Hao Wang, Tao Song, Zhengui Xue, Ruhui Ma, and Haibing Guan. Fedala: Adaptive local aggregation for personalized federated learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 11237–11244, 2023.
- [Zhao and Zhang, 2020] Shuping Zhao and Bob Zhang. Joint constrained least-square regression with deep convolutional feature for palmprint recognition. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(1):511–522, 2020.

- [Zhao *et al.*, 2019] Shuping Zhao, Bob Zhang, and CL Philip Chen. Joint deep convolutional feature representation for hyperspectral palmprint recognition. *Information Sciences*, 489:167–181, 2019.
- [Zhong and Zhu, 2019] Dexing Zhong and Jinsong Zhu. Centralized large margin cosine loss for open-set deep palmprint recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(6):1559–1568, 2019.