A. Experimental Details

We use 1 NVIDIA RTX A5000 with 24GB RAM to run the experiments. We use PyTorch [3] to implement our algorithm. For the baseline FedAvg, the total communication round is set to 50. For FGL [6], we generate 3500 images per class per domain. For the optimization of the classification model, we use SGD with momentum as the optimizer, where the learning rate is set to 0.01 and the momentum is 0.9. The optimization epoch is set to 50. The training image resolution is set to 512×512 for all datasets.

For FedD3 [4], we adopt Kernel Inducing Points (KIP) to distill the original dataset into 1 image per class per domain and transmit them to the central server. For DENSE [5], we first finetune the pretrained ResNet-18 [1] at each client and then optimize a Generator to conduct model distillation at central server. The hyperparameters used in these methods are following their original papers. For `FedMLA`, we use Adam optimizer to optimize the concept vectors. The learning rate is set to 0.1 and beta is set to (0.9, 0.999). The total training epochs is set to 30. We adopt the Pseudo Numerical Diffusion Model (PNDM) [2] in the Latent Diffusion Model. The perturbation intensity for domain concept vector σ_μ is set to 0.1 for all dataset. More dataset specific hyperparameters are provided in Table 1.

B. Synthetic Image Visualization

We provide synthetic images for all benchmarks in the following figures, where we observe that the synthetic images generally follow the distribution and characteristics of the original training datasets at each client. Besides, the visual quality of the generated images, e.g., the detailed features of the objects, is also promising.

C. Communication and Computation Costs

FedBiP introduces optimization exclusively during the concept-level personalization phase, wherein only the concept vectors are updated, while the weights of the Latent Diffusion Model (LDM) remain unchanged. Importantly, no additional optimization steps are performed during instance-level personalization or server-side image generation. This design highlights the computational efficiency of the proposed approach, making it particularly suitable for federated learning (FL) applications involving a large number of clients with limited computational power.

Moreover, during the one-shot upload process, only the concept vectors and latent vectors are transmitted. The dimensions of the concept vectors remain fixed, whereas the size of the latent vectors varies based on the method's scale (e.g., S, M, or L) and can be adjusted to align with the capacity of different FL systems.

References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 1

[2] Luping Liu, Yi Ren, Zhijie Lin, and Zhou Zhao. Pseudo numerical methods for diffusion models on manifolds. *arXiv preprint arXiv:2202.09778*, 2022. 1

[3] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019. 1

[4] Rui Song, Dai Liu, Dave Zhenyu Chen, Andreas Festag, Carsten Trinitis, Martin Schulz, and Alois Knoll. Federated learning via decentralized dataset distillation in resource-constrained edge environments. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–10. IEEE, 2023. 1

[5] Jie Zhang, Chen Chen, Bo Li, Lingjuan Lyu, Shuang Wu, Shouhong Ding, Chunhua Shen, and Chao Wu. Dense: Data-free one-shot federated learning. *Advances in Neural Information Processing Systems*, 35:21414–21428, 2022. 1

[6] Jie Zhang, Xiaohua Qi, and Bo Zhao. Federated generative learning with foundation models. *arXiv preprint arXiv:2306.16064*, 2023. 1

Table 1. Detailed hyperparameters for each dataset. The highlighted words ([STY]) in the textual prompt will be replaced by the domain concept vectors. The [CLS] will be replaced by the class concept vectors.

Dataset	prompt	n_s	n_c	C	Class Names
Derma MNIST	A dermatoscopic image of a [CLS], a type of pigmented skin lesions.	2	4	10	intraepithelial carcinoma, basal cell carcinoma, benign keratosis, dermatofibroma, melanoma, melanocytic nevi, vascular skin
UCM	A centered satellite photo of [CLS].	3	3	21	agricultural, dense residential, medium residential, sparse residential, parking lot, buildings, harbor, mobile homepark, storage tanks, freeway, intersection, overpass, golf course, baseball diamond, runway, tenniscourt, beach, forest, river, chaparral, airplane
Domain Net	A [STY] of [CLS].	1	1	10	airplane, clock, axe, basketball, bicycle, bird, strawberry, flower, pizza, bracelet
Office Home	A [STY] of [CLS].	1	1	20	Marker, Spoon, Pencil, Speaker, Toys, Fan, Hammer, Notebook, Telephone, Sink, Chair, Fork, Kettle, Bucket, Knives, Monitor, Mop, Oven, Pen, Couch
PACS	A [STY] of [CLS].	1	1	7	dog, elephant, giraffe, guitar, horse, house, person

Figure 1. Synthetic Images for DomainNet benchmark.

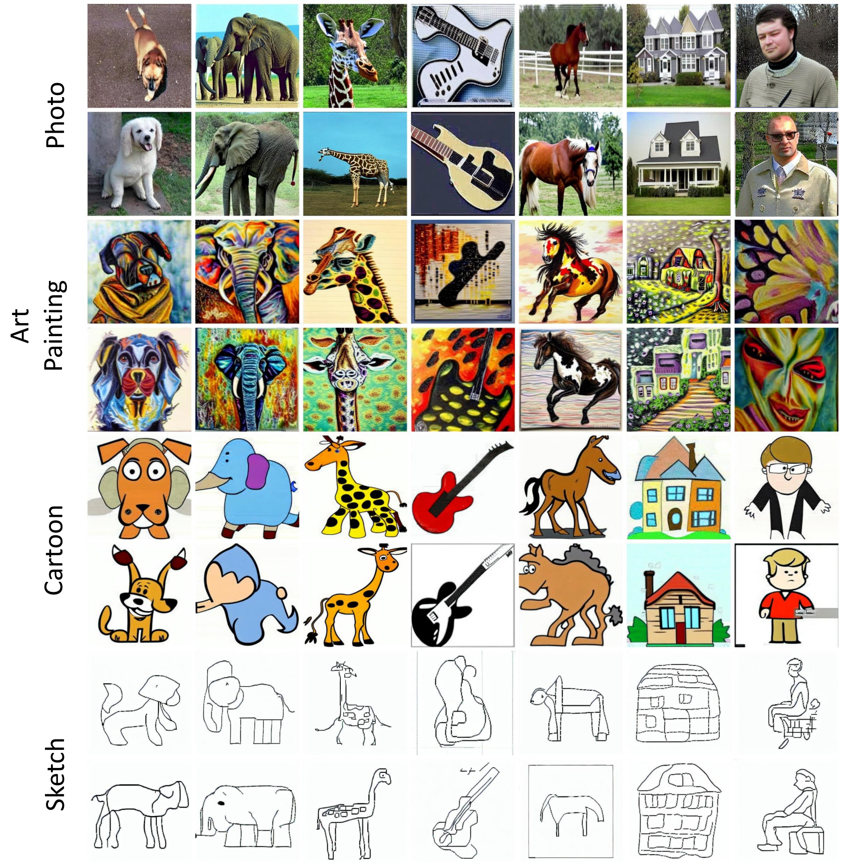

Figure 2. Synthetic Images for PACS benchmark.

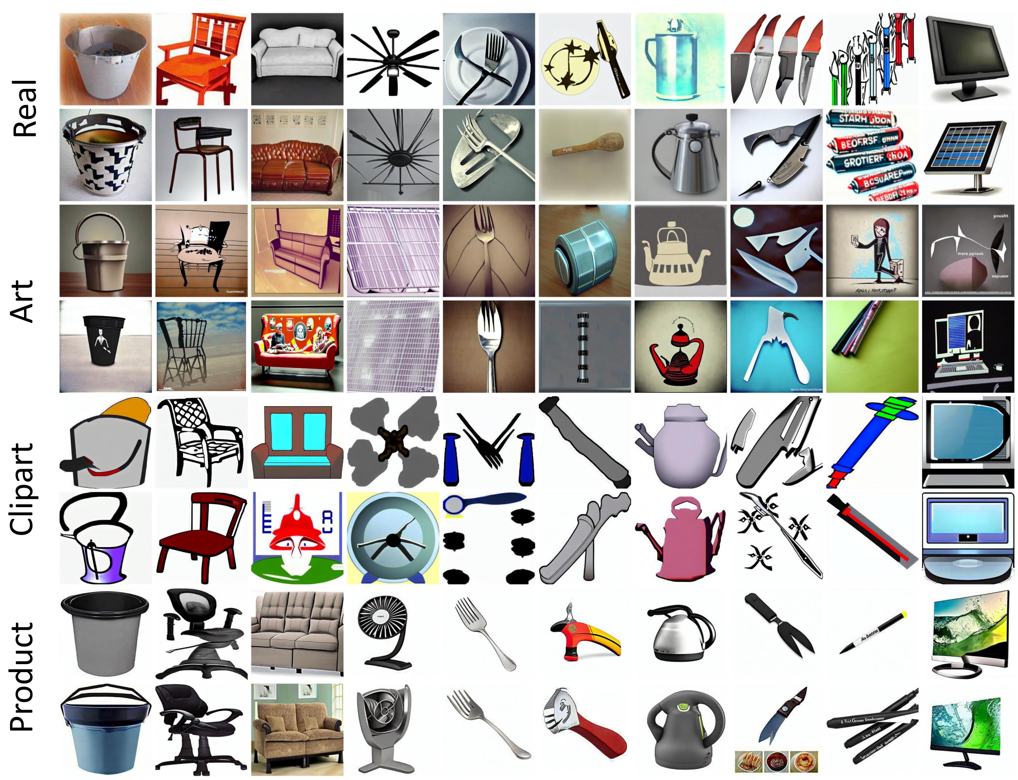

Figure 3. Synthetic Images for OfficeHome benchmark.

Figure 4. Synthetic Images for UCM benchmark.

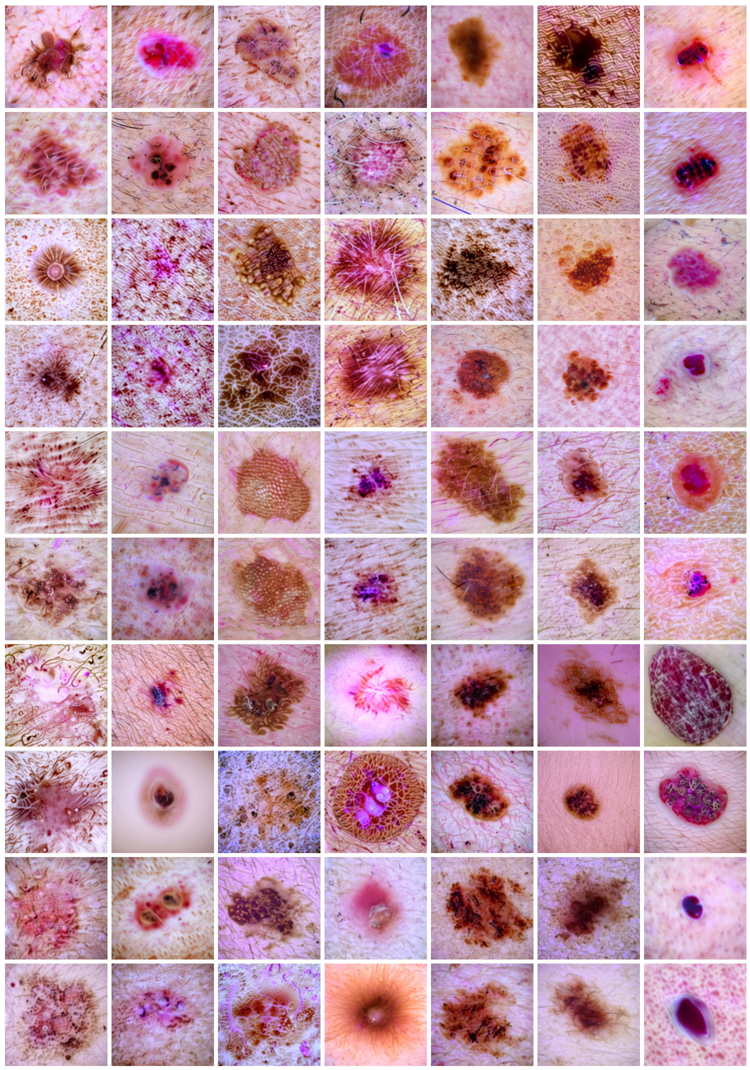

Figure 5. Synthetic Images for DermaMNIST benchmark.