

# DP-GENG : Differentially Private Dataset Distillation Guided by DP-Generated Data

Shuo Shi<sup>1,2</sup>, Jinghui Zhang<sup>3</sup>, Shijie Jiang<sup>4</sup>, Chunyi Zhou<sup>1</sup>,  
Yuyuan Li<sup>5</sup>, Mengying Zhu<sup>1</sup>, Yangyang Wu<sup>1</sup>, Tianyu Du<sup>1,2\*</sup>

<sup>1</sup>Zhejiang University

<sup>2</sup>Ningbo Global Innovation Center, Zhejiang University

<sup>3</sup>University of California, Los Angeles

<sup>4</sup>Hangzhou Normal University

<sup>5</sup>Hangzhou Dianzi University

{shuoshi, zjradty}@zju.edu.cn

## Abstract

Dataset distillation (DD) compresses large datasets into smaller ones while preserving the performance of models trained on them. Although DD is often assumed to enhance data privacy by aggregating over individual examples, recent studies reveal that standard DD can still leak sensitive information from the original dataset due to the lack of formal privacy guarantees. Existing differentially private (DP)-DD methods attempt to mitigate this risk by injecting noise into the distillation process. However, they often fail to fully leverage the original dataset, resulting in degraded realism and utility. This paper introduces DP-GENG, a novel framework that addresses the key limitations of current DP-DD by leveraging DP-generated data. Specifically, DP-GENG initializes the distilled dataset with DP-generated data to enhance realism. Then, generated data refines the DP-feature matching technique to distill the original dataset under a small privacy budget, and trains an expert model to align the class distribution. Furthermore, We provide a privacy budget allocation strategy with theoretical guarantees. Extensive experiments demonstrate that DP-GENG significantly outperforms state-of-the-art methods in both utility and privacy robustness.

**CODE** — <https://github.com/shuoshiss/DP-GENG>

**Extended version** — <https://arxiv.org/abs/2511.09876>

## 1 Introduction

Dataset distillation (DD) (Wang et al. 2018) aims to compress large datasets into smaller ones while preserving the utility of models trained on them. This technique offers compelling advantages: it enhances training efficiency, reduces storage demands (Yu, Liu, and Wang 2024), and can improve data privacy when working with sensitive data (Dong, Zhao, and Lyu 2022; Tong et al. 2025, 2024). Recent advances (Guo et al. 2024; Sun et al. 2024b; Wang et al. 2025) have shown that models trained on distilled datasets can achieve accuracy comparable to those trained on full datasets—even when using orders of magnitude fewer training examples.

\*Corresponding Author

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

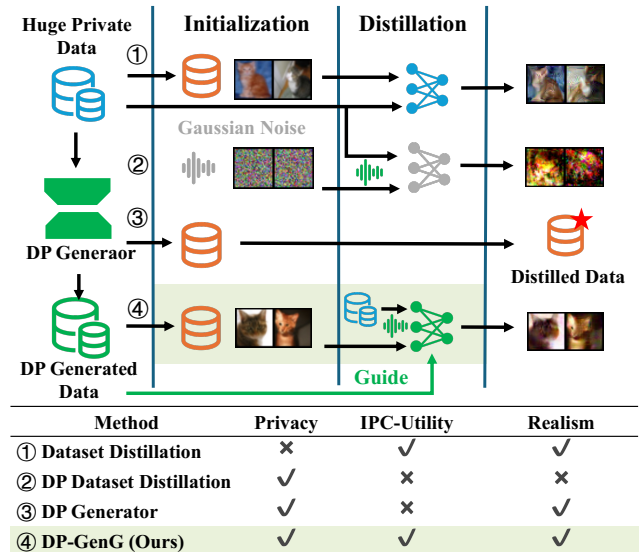


Figure 1: Comparison of privacy, utility, and realism across dataset distillation methods.

Despite the promising performance of DD, a critical privacy issue remains: although the distilled dataset is assumed to inherently protect data privacy due to its compressed nature, recent studies (Carlini, Feldman, and Nasr 2022; Chen, Kerkouche, and Fritz 2022) have shown that standard distillation techniques offer no formal privacy guarantees and can still leak sensitive information. This limitation raises significant concerns, particularly when applied to sensitive datasets—such as medical records—where even marginal leakage of private details could violate privacy regulations (e.g., GDPR (Regulation 2016)) or expose individuals to harm (Karale 2021; Lu, Tong, and Ye 2025; Lan et al. 2025).

To address this issue, existing methods (e.g., PSG (Chen, Kerkouche, and Fritz 2022) and NDPDC (Zheng and Li 2024)) propose differentially private dataset distillation (DP-DD) techniques. The key idea is to inject Gaussian noise into the distillation process, allowing the creation of a distilled dataset from the private dataset while ensuring formal differ-

ential privacy (DP) guarantees. While these methods limit an adversary’s ability to infer the presence of individual data points, they suffer from key limitations that result in suboptimal performance, as shown in Figure 1. **Limitation 1:** They fail to preserve the realism of each distilled example due to the lack of direct access to natural data. Realism quantifies how visually and semantically consistent an example is with its corresponding label. A low realism score - often stemming from poor initialization or noisy optimization signals - can severely compromise the utility of the distilled dataset (Shao et al. 2024). **Limitation 2:** Given a limited privacy budget, these methods must add large amounts of noise, which further degrades both the quality and utility of the resulting dataset.

Previous methods primarily enforce DP by injecting noise directly into the distillation process, which limits their ability to fully exploit the information in the private dataset. Motivated by recent advances in DP-synthetic data generation, which produces synthetic data that closely resemble the private data distribution while ensuring formal DP guarantees, we propose DP-GENG (Differentially Priate Generator-Guided Distillation). DP-GENG is a novel DP-DD pipeline that bridges DP-generated data with DD to address key limitations in existing DP-DD frameworks. Specifically, this work aims to address the following research questions: **(RQ1)** *How can we leverage DP-generated data to guide the distillation process and address key limitations?* **(RQ2)** *How can we increase the utility of distilled datasets under limited privacy budgets?*

DP-GENG begins by using DP data generators to produce a large volume of synthetic data that closely resembles the original private dataset. Then, it selects a representative subset of DP-generated data to initialize the distillation process, which substantially promotes the realism of distilled examples (**solution to L1**). In the subsequent phase, DP-GENG introduces a novel DP feature-matching technique, which aligns the small distilled dataset with the original dataset in the feature space. Unlike previous methods (Zhao et al. 2023; Zhang et al. 2023, 2024) that train multiple feature extractors directly on the private dataset for meaningful guidance, DP-GENG constructs feature extractors using the DP-generated data. Due to the post-processing property of DP (Dwork, Roth et al. 2014), this design incurs no additional privacy cost, thereby mitigating the impact of budget limitations (**solution to L2**). Moreover, we identify a critical issue: directly applying DP to DD can degrade performance due to the noises introduced in the matching process, which may cause the distilled examples to drift away from their intended class representation. To address this, DP-GENG introduces a novel DP expert model that acts as a calibrator to dynamically align each distilled example with samples from the same class during distillation.

To further ensure high-quality distillation within a limited privacy budget, DP-GENG designs a tailored privacy budget allocation algorithm that determines budget consumption across different DP components (e.g., synthetic data generation, feature matching, expert guidance) and provides a theoretical analysis of the overall privacy guarantees ensured by DP-GENG. Extensive experiments demonstrate that DP-GENG significantly outperforms existing DP-DD methods in

both utility and privacy trade-offs. Our contributions can be summarized as follows:

- We propose DP-GENG, a novel framework that bridges DP-generated data with differentially private dataset distillation. DP-GENG identifies key limitations in existing DP-DD methods and leverages tailored mechanisms, guided by DP-generated data, to address them.
- We design a privacy budget allocation strategy across DP components and provide a theoretical analysis of the overall privacy guarantees ensured by DP-GENG.
- Our evaluations show DP-GENG not only outperforms prior DPDD methods by 11.6% but is also highly robust to membership inference attacks.

## 2 Related Work

**Dataset Distillation (DD).** DD aims to compress large datasets into smaller synthetic ones while preserving their utility for model training. Early works (Wang et al. 2018; Zhao, Mopuri, and Bilen 2020; Zhao and Bilen 2023) pioneer gradient matching and distribution matching approaches. Recent advances have further improved DD through more advanced techniques, such as trajectory matching (Cazenavette et al. 2022; Guo et al. 2024), and parameterization (Liu et al. 2022; Yuan et al. 2024).

While DD is often assumed to protect data privacy (Li et al. 2020; Dong, Zhao, and Lyu 2022), recent studies (Li et al. 2024b; Zhao and Zhang 2025a) show that standard DD methods remain vulnerable to membership inference attacks (Carlini et al. 2022). To address this, many works have proposed integrating DP into DD to develop DP-DD methods with formal privacy guarantees, including PSG (Chen, Kerkouche, and Fritz 2022) and NDPDC (Zheng and Li 2024). However, these methods simply inject noise into the distillation process and fail to fully exploit the information in the private dataset, leading to suboptimal results.

**Differential Privacy (DP).** DP (Dwork et al. 2006) provides a theoretical framework for privacy protection by ensuring that the presence or absence of any individual record in a dataset has minimal impact on an algorithm’s output (See basic definition of DP in Appendix (Shi et al. 2025)). This is achieved by injecting calibrated noise, which prevents adversaries from inferring sensitive information about individual data while maintaining statistical utility (Dwork et al. 2006; Patwa et al. 2023; Sun et al. 2024a; Farayola, Olorunfemi, and Shoetan 2024). Shokri et al. (2017) show that applying DP to a learning task can reduce the success rate of privacy attacks. Besides, Jayaraman and Evans (2019) evaluate the effectiveness of  $(\epsilon, \sigma)$ -DP and its variants in neural networks by using membership inference attacks (Shokri et al. 2017).

Prior to our work, DP-DD methods are based on Rényi Differential Privacy (RDP) (Mironov 2017). In this work, we adopt  $f$ -DP (Dong, Roth, and Su 2022) to achieve improved performance, which offers lossless privacy accounting through a hypothesis testing framework. More specifically, consider testing  $H_0$ : data  $\sim P$  vs.  $H_1$ : data  $\sim Q$  with a rejection rule  $\phi \in [0, 1]$ . The type I error is  $\alpha_\phi = \mathbb{E}_P[\phi]$ , and the type II error is  $\beta_\phi = 1 - \mathbb{E}_Q[\phi]$ . The trade-off function  $T(P, Q)(\alpha)$  in  $f$ -DP gives the minimal type II error for type

I error at level  $\alpha$ :  $T(P, Q)(\alpha) = \inf_{\phi} \{\beta_{\phi} : \alpha_{\phi} \leq \alpha\}$ . Specifically, this work utilizes  $\mu$ -Gaussian Differential Privacy ( $\mu$ -GDP), an instance of f-DP where the trade-off function is given by  $f = G_{\mu}$ , with  $G_{\mu}(x) = \Phi(\Phi^{-1}(1-x) - \mu)$  and  $\Phi$  denoting the cumulative distribution function of the standard normal distribution (see Appendix B for details).

**DP-synthetic Data Generation.** The practical demand for privacy-preserving synthetic data drives the development of DP-synthetic data generation methods. They aim to generate a dataset (called DP-generated data) that retains the properties of the private dataset while protecting individual privacy (Jia et al. 2025). For example, Ghalebikesabi et al. (2023) fine-tunes a pre-trained generator on the private dataset with DP guarantees, then uses the fine-tuned generator to produce DP-generated data. In contrast, Lin et al. (2023) directly uses a pre-trained generator to create DP-generated data, injecting noise during the dataset refinement process.

### 3 Methodology

In this part, we present our framework, DP-GENG, which integrates DP-generated data to improve the performance of DP-DD. Following (Chen, Kerkouche, and Fritz 2022; Zheng and Li 2024), we focus on DP-DD in the image domain. Section 3.1 describes the techniques used to generate synthetic datasets under DP guarantees. Section 3.2 explains how DP-generated data is used to guide each component of the DD process. Section 3.3 introduces an expert model designed to mitigate issues introduced by DP noise. Finally, Section 3.4 details the budget allocation and analyzes the overall privacy guarantees ensured by DP-GENG.

#### 3.1 DP Data Generation

Inspired by recent advances in DP image synthesis, we aim to generate a synthetic dataset  $\hat{\mathcal{T}}$  that closely resembles the private dataset  $\mathcal{T}$  while providing formal privacy guarantees. Existing DP image synthesis approaches differ in how they ensure privacy and can be categorized into input-level (Harder, Adamczewski, and Park 2021), model-level (Ghalebikesabi et al. 2023; Liu et al. 2023) and output-level (Lin et al. 2023) methods, depending on the stage at which DP noise is applied. Among them, the most widely adopted methods involve pre-training generators on large public datasets and subsequently fine-tuning them on the private dataset  $\mathcal{T}$  using DP mechanisms (e.g., PrivImage (Li et al. 2024a)), or injecting noise directly during the synthesis and refinement phases (e.g., PE (Lin et al. 2023)). The two approaches are the primary focus of this paper. Notably, these approaches rely on the Gaussian mechanism, where Gaussian noise  $\xi$  is introduced either during the fine-tuning phase of DP data generators (Li et al. 2024a) or in the voting results of evolutionary algorithms (Lin et al. 2023). For such methods, adding Gaussian noise with a standard deviation  $\sigma_G = 1/\mu_G$  ensures a privacy guarantee of  $\mu_G$ -GDP, as formalized below:

**Lemma 1 (Gaussian Mechanism to GDP (Dong, Roth, and Su 2022)).** *Define the Gaussian mechanism that operates on a statistic  $\theta$  as  $\mathcal{M}(D) = \theta(D) + \xi$ , where  $\xi \sim \mathcal{N}(0, \text{sens}(\theta)^2/\mu^2)$ . Then,  $\mathcal{M}$  is  $\mu$ -GDP.*

DP image synthesis approaches can produce a large number of synthetic samples  $\hat{\mathcal{T}}$  that capture diverse characteristics of the original private dataset. Moreover, due to the post-processing property of DP (Dwork, Roth et al. 2014), the DP-generated data  $\hat{\mathcal{T}}$  will inherit the privacy guarantees of the synthesis process. In other words,  $\hat{\mathcal{T}}$  can be freely utilized in downstream computations (distillation process), without incurring any additional privacy cost.

**Theorem 1 (Post-processing property (Dwork, Roth et al. 2014)).** *If  $\mathcal{M}$  satisfies  $(\epsilon, \delta)$ -DP,  $\mathcal{F} \circ \mathcal{M}$  will satisfy  $(\epsilon, \delta)$ -DP for any private data independent function  $\mathcal{F}$  with  $\circ$  denoting the composition operator.*

Motivated by this, **the key idea is to construct a large synthetic dataset that can replace the private dataset during certain parts of the distillation process.** In contrast to prior methods that enforce DP solely by injecting noise during the optimization, the use of DP-generated data offers a novel perspective on achieving privacy guarantees, allowing us to fully leverage the knowledge contained in the private dataset without incurring large privacy costs while improving the realism of the distilled dataset.

#### 3.2 DP Feature Matching

Feature matching (FM) is an effective method for DD, where the distilled dataset is optimized to match the feature distributions of the full dataset. We adopt FM as the core distillation algorithm within DP-DD due to its generalizability. Moreover, we incorporate DP-generated data in both the initialization and matching phases to address identified limitations.

**Initialization with DP-generated Data.** A recent study shows that DD methods achieve optimal utility when distilled examples are initialized with samples from the private dataset (Shao et al. 2024; Zhao and Chen 2023), which is attributed to the enhanced realism provided by such initialization. Unlike these approaches, as well as previous DP-DD methods that rely on Gaussian noise initialization (Chen, Kerkouche, and Fritz 2022; Zheng and Li 2024), we propose using DP-generated data to initialize the distilled dataset. Given that the size of the distilled dataset is controlled by the number of images per class (IPC), we construct the initialized distilled set  $\mathcal{S}_{\text{init}}$  by applying a sampling strategy (e.g., selection by performing k-means clustering on features) over the DP-generated data  $\hat{\mathcal{T}}$ :  $\mathcal{S}_{\text{init}} = \text{Sample}(\hat{\mathcal{T}}, \text{IPC})$ . Moreover, we employ a parameterization technique (Kim et al. 2022), which embeds multiple DP synthetic images within a single image, to maximize the utilization of the private dataset. Notably, the use of DP-generated data for initialization directly addresses limitation **(L1)** by preserving the realism of the distilled dataset, which in turn leads to enhanced utility-privacy tradeoff.

**Feature Extractor from DP-generated Data.** The feature extractor in FM captures image representations to facilitate the alignment of feature distributions. The success of FM relies on multiple trained feature extractors (Zhao et al. 2023; Zhang et al. 2024; Zhao 2024). Existing methods either use private dataset directly to train the feature extractors, which violates DP guarantees, or rely on randomly initialized feature extractors, which results in poor utility. A feasible al-

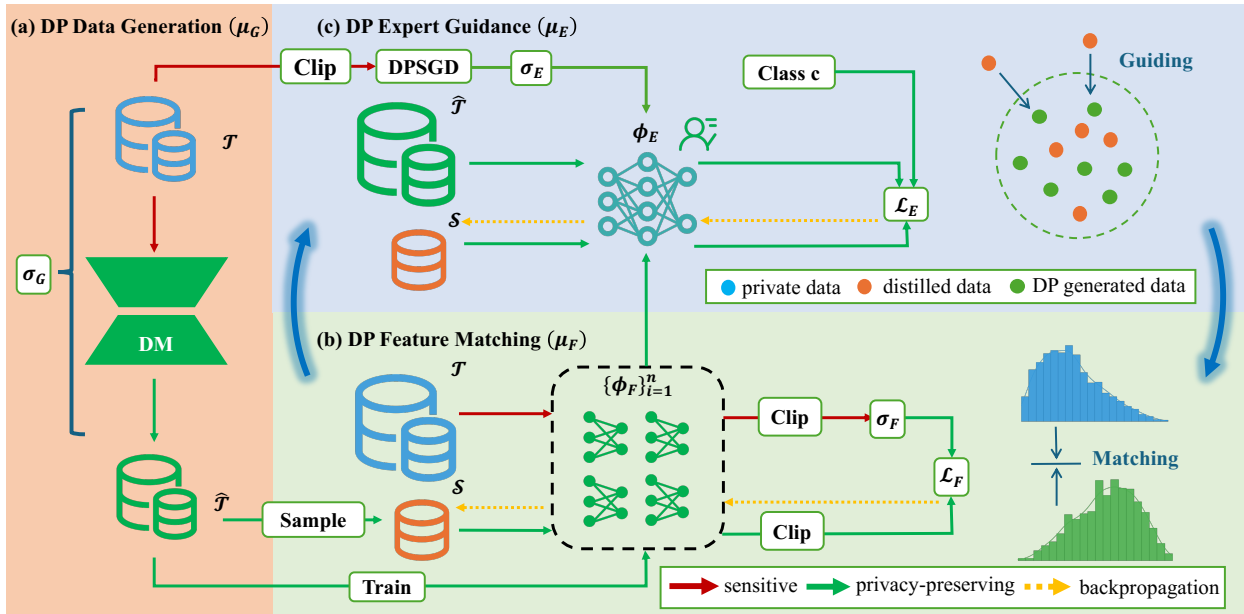


Figure 2: The overall framework. It fully leverages DP-generated data throughout the distillation process to enhance the performance of the distilled dataset. The blue, green and orange datasets represent the original private dataset, its DP-generated version and DP-distilled dataset.

ternative is to obtain DP-guaranteed feature extractors using DP-SGD (Abadi et al. 2016); however, this approach is often impractical under limited privacy budgets. Specifically, achieving robust feature extraction requires training multiple feature extractors  $\{\phi_F^n\}_{n=1}^N$  in parallel over multiple iterations. Applying DP-SGD in this setting introduces substantial noise due to the limited privacy budgets allocated to each training run, which ultimately degrades the quality of the distilled dataset. In this work, we propose training feature extractors using DP-generated data, which retains the ability to capture features of the private dataset while effectively addressing the budget limitation (L2). According to Theorem 1, models trained on DP-generated data inherit its privacy guarantees without incurring additional privacy cost.

**Matching via Injected DP Noise.** Given trained feature extractors, the matching phase aligns the feature distribution of the distilled dataset with that of a full dataset. This full dataset can be either DP-generated data or the original private dataset. The only difference is that, when using the original private dataset, DP must be enforced by injecting noise during the matching process. Empirically, we find that using the original private dataset yields better performance under the same privacy budget, despite the added noise. We attribute this to the fact that the distilled dataset is initialized with DP-generated data, which already captures prior knowledge from that source. As a result, the original private dataset provides a more informative guidance signal during matching. Specifically, we inject Gaussian noise with standard deviation  $\sigma_F$  during the matching process, as shown below:

$$\bar{\phi}_F(\mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{\mathbf{x} \in \mathcal{D}} \text{CLIP}(\phi_F(\mathbf{x}), C), \quad (1)$$

$$\mathcal{L}_F = \|\bar{\phi}_F(\mathcal{T}) + \mathcal{N}(0, \sigma_F^2 C^2 I) - \bar{\phi}_F(\mathcal{S})\|^2. \quad (2)$$

where  $\text{CLIP}(\phi_F(\mathbf{x}^*), C) = \min(1, \frac{C}{\|\phi_F(\mathbf{x}^*)\|}) \phi_F(\mathbf{x}^*)$  represents a feature clipping mechanism designed to regulate the sensitivity of the L2 norm of the averaged feature vector.

Unlike prior work (Chen, Kerkouche, and Fritz 2022; Zheng and Li 2024) that relies on Rényi Differential Privacy (Mironov 2017), we adopt GDP (Dong, Roth, and Su 2022). Due to mini-batch processing, by applying the **Sub-sampling Theorem** of GDP in Appendix F, we achieve noise reduction proportional to the sampling probability  $p$ , while preserving the same privacy guarantees. The resulting privacy parameter is given by:  $\mu_F = p\sqrt{T(e^{1/\sigma_F^2} - 1)}$ .

### 3.3 DP Expert Guidance

After initialization, examples in the distilled dataset largely retain their visual appearance throughout the matching process. However, their feature representations are continuously optimized to capture the characteristics of the full dataset. Due to the effects of DP noise introduced in DP-DD algorithms, we observe that some distilled examples may undergo significant shifts in their feature representations. As illustrated in Appendix J (Shi et al. 2025), these shifts can cause an example to deviate from its original class, degrading the utility of the distilled dataset. To address this, we introduce an expert model to regularize the matching process and keep features of distilled examples aligned with their class distributions.

**Expert Model.** We employ an expert model to learn class-wise feature representations, which enables the calibration of misaligned distilled examples affected by DP noise. To train this expert, we first perform standard training on DP-generated data, followed by fine-tuning with DP-SGD on the

original private dataset with  $\mu_E$ . See Appendix C for details. **DP-Generated Data Guided Alignment.** We aim to align distilled examples with their class representations by using the expert model as a regularizer. For each distilled example, we sample reference data points  $\mathcal{R}^{y^S}$  from its class  $y^S$  and encourage the example to follow the same feature distribution as these references (Ma and Yang 2024). To preserve privacy, we use samples from the DP-generated data as references without additional privacy cost. Since the expert model is trained to accurately capture class-wise distinctions, the alignment task becomes equivalent to maximizing the similarity between each distilled example and its associated references in the expert model’s label space. For each distilled example, we aggregate the predicted logits of its references and apply two forms of supervision: (1) KL divergence between the distilled example’s logits and the aggregated soft labels to guide distributional alignment, and (2) hard-label regularization to preserve class identity. We illustrate the loss at each iteration below, where  $\mathcal{R}^{y^S} \in \hat{\mathcal{T}}$  represents references that are randomly sampled from class  $y^S$ :

$$c_y = \frac{1}{|\mathcal{R}^{y^S}|} \sum_{\mathbf{x}^{\mathcal{R}} \in \mathcal{R}^{y^S}} \phi_E(\mathbf{x}^{\mathcal{R}}), \quad (3)$$

$$\mathcal{L}_E = \frac{1}{|\mathcal{S}|} \sum_{(\mathbf{x}^S, \mathbf{y}^S) \in \mathcal{S}} \left[ l_{\text{CE}}(\phi_E(\mathbf{x}^S), \mathbf{y}^S) + \lambda \text{D}_{\text{KL}}(\phi_E(\mathbf{x}^S) \| c_{y^S}) \right]. \quad (4)$$

### 3.4 Overall Privacy Analysis

DP-GENG involves three components that jointly consume the privacy budget: DP data generation  $\mu_G$ , feature matching  $\mu_F$ , and expert model training  $\mu_E$ . In this part, we first analyze the total privacy cost of DP-GENG. We then present a practical strategy for selecting optimal parameters to achieve a strong privacy-utility trade-off. The algorithm of DP-GENG is shown in Algorithm 1 in Appendix C (Shi et al. 2025).

**Privacy Cost Calculation for the Entire Process.** The privacy parameters  $\mu_G$ ,  $\mu_F$ , and  $\mu_E$ , which are controlled by the Gaussian noise, determine the total privacy cost. Based on Lemma 2, we compute the overall privacy parameter  $\mu_{\text{total}}$  for the entire DP-GENG workflow as below:

**Lemma 2 (GDP Composition (Dong, Roth, and Su 2022)).** *The  $n$ -fold composition of  $\mu_i$ -GDP mechanisms is  $\sqrt{\mu_1^2 + \dots + \mu_n^2}$ -GDP.*

Next, we need to convert the  $\mu$ -GDP to  $(\epsilon, \delta)$ -DP. We can easily connect GDP and DP by Lemma 3.

**Lemma 3 (GDP to DP Conversion (Dong, Roth, and Su 2022)).** *A mechanism is  $\mu$ -GDP if and only if it is  $(\epsilon, \delta(\epsilon))$ -DP for all  $\epsilon \geq 0$ , where*

$$\delta(\epsilon) = \Phi\left(-\frac{\epsilon}{\mu} + \frac{\mu}{2}\right) - e^{\epsilon} \Phi\left(-\frac{\epsilon}{\mu} - \frac{\mu}{2}\right)$$

where  $\Phi$  denotes the standard normal CDF.

**Budget Allocation Strategy.** The overall privacy cost of DP-GENG depends on the amount of noise injected into its components: DP data generation ( $\sigma_G$ ), DP feature matching ( $\sigma_F$ ), and DP expert model training ( $\sigma_E$ ). The corresponding

GDP parameters are  $\mu_G$ ,  $\mu_F$ , and  $\mu_E$ . We strategically allocate the privacy budget by prioritizing the noise levels for DP data generation and expert model training ( $\sigma_G$  and  $\sigma_E$ ), since their performance can be independently evaluated. Specifically, we employ binary search to determine the appropriate noise multiplier  $\sigma_G$  for DP data generation, aiming to achieve a target FID score. Similarly, we adjust  $\sigma_E$  for training the expert model to reach a desired accuracy (a detailed explanation of the target FID score and accuracy can be found in Appendix D). Once the noise multipliers  $\sigma_G$  and  $\sigma_E$  are set based on these utility goals, we compute the required noise level  $\sigma_F$  for DP feature matching using Lemma 2, ensuring that the overall privacy budget  $(\epsilon_{\text{total}}, \delta_{\text{total}})$  is satisfied. In other words, we allocate the privacy budget to each component according to a utility-driven criterion defined by the target FID of the DP-generated data and the target accuracy of the expert model. Our strategy ensures that the total privacy expenditure remains within the specified budget, as detailed in the following theorem, proved in Appendix F.

**Theorem 2 (DP-GENG Privacy Budget Allocation).** *In the DP-GENG, each component involves differential privacy guarantees. Given the total privacy parameters  $\epsilon_{\text{total}}$  and  $\delta_{\text{total}}$ , with  $\mu_G$  for generation and  $\mu_E$  for expert guidance stages, we can derive the noise parameter  $\sigma_F$  for feature matching with batch sampling probability  $p = \frac{B}{n}$  over  $T$  iterations as follows:*

$$\sigma_F = \sqrt{\ln \frac{T \cdot p^2}{\mu_{\text{total}}^2 - \mu_G^2 - \mu_E^2} + 1}^{-1}.$$

## 4 Experiment

### 4.1 Experiment Setup

**Datasets and Models.** Unlike previous works on DP-DD (Chen, Kerkouche, and Fritz 2022; Zheng and Li 2024), which focus on simple datasets such as MNIST (LeCun et al. 1998) and FashionMNIST (Xiao, Rasul, and Vollgraf 2017), we conduct experiments on the CIFAR-10 (Krizhevsky, Hinton et al. 2009), CIFAR-100 (Krizhevsky, Hinton et al. 2009), and CelebA (Liu et al. 2015) datasets. These datasets present significant challenges for DP-DD, especially under a limited privacy budget. Following previous studies (Guo et al. 2024; Zheng and Li 2024), we use ConvNet (Sagun et al. 2017) as the default backbone architecture to evaluate the utility of the distilled dataset. Additionally, we present results for other models (e.g., ResNet (He et al. 2016)) in Appendix H (Shi et al. 2025), which benefit from the prior knowledge in DP-generated data to enhance cross-architecture generalizability. **Baselines.** We compare our proposed method, DP-GENG, with state-of-the-art DP-DD algorithms and distilled datasets directly generated from DP data generators. For comprehensive evaluation, all experiments are conducted using three different random seeds, and we report both the mean and variance of the results.

- DP-DD Methods: (i) Gradient matching based method: PSG (Chen, Kerkouche, and Fritz 2022); (ii) Distribution matching based method: NDPDC (Zheng and Li 2024).

	CIFAR-10			CIFAR-100			CelebA		
IPC	1	10	50	1	10	50	1	10	50
$\epsilon = 1$									
DP-MERF	14.6 $\pm$ 0.7	19.4 $\pm$ 0.3	21.0 $\pm$ 0.4	2.3 $\pm$ 0.1	3.4 $\pm$ 0.0	3.6 $\pm$ 0.1	53.6 $\pm$ 0.4	64.6 $\pm$ 0.8	68.3 $\pm$ 0.3
PE	15.2 $\pm$ 0.2	25.0 $\pm$ 1.0	37.7 $\pm$ 1.0	3.0 $\pm$ 0.3	7.4 $\pm$ 0.3	11.1 $\pm$ 0.2	54.8 $\pm$ 0.3	65.1 $\pm$ 0.3	70.6 $\pm$ 0.3
PrivImage	14.7 $\pm$ 0.2	21.2 $\pm$ 0.3	38.0 $\pm$ 0.2	1.2 $\pm$ 0.1	2.0 $\pm$ 0.1	3.8 $\pm$ 0.1	54.5 $\pm$ 1.3	64.2 $\pm$ 0.7	72.1 $\pm$ 0.6
PSG	26.9 $\pm$ 0.7	33.7 $\pm$ 0.2	35.9 $\pm$ 0.6	4.9 $\pm$ 0.3	8.3 $\pm$ 0.3	10.3 $\pm$ 0.1	66.8 $\pm$ 3.8	72.8 $\pm$ 2.1	76.2 $\pm$ 2.3
NDPDC	26.1 $\pm$ 0.4	39.8 $\pm$ 0.2	42.6 $\pm$ 0.6	7.8 $\pm$ 0.2	10.9 $\pm$ 0.1	11.5 $\pm$ 0.3	66.0 $\pm$ 1.7	77.7 $\pm$ 0.5	80.4 $\pm$ 0.6
<b>DP-GENG (Ours)</b>	<b>29.8<math>\pm</math>0.5</b>	<b>53.5<math>\pm</math>0.3</b>	<b>56.9<math>\pm</math>0.3</b>	<b>15.8<math>\pm</math>0.3</b>	<b>21.1<math>\pm</math>0.4</b>	<b>25.9<math>\pm</math>0.2</b>	<b>67.8<math>\pm</math>0.7</b>	<b>78.7<math>\pm</math>0.3</b>	<b>82.1<math>\pm</math>0.1</b>
$\epsilon = 10$									
DP-MERF	14.9 $\pm$ 0.3	21.7 $\pm$ 0.3	22.9 $\pm$ 0.5	2.8 $\pm$ 0.3	3.4 $\pm$ 0.2	3.8 $\pm$ 0.2	54.1 $\pm$ 0.6	68.5 $\pm$ 0.2	72.2 $\pm$ 0.7
PE	16.0 $\pm$ 0.4	29.8 $\pm$ 0.4	42.1 $\pm$ 0.3	3.4 $\pm$ 0.1	10.0 $\pm$ 0.2	15.1 $\pm$ 0.2	56.1 $\pm$ 0.4	67.5 $\pm$ 0.5	75.3 $\pm$ 0.4
PrivImage	15.2 $\pm$ 0.4	27.7 $\pm$ 0.5	39.6 $\pm$ 0.5	1.9 $\pm$ 0.1	2.8 $\pm$ 0.2	4.3 $\pm$ 0.2	55.2 $\pm$ 0.4	65.7 $\pm$ 0.2	76.8 $\pm$ 0.3
PSG	27.9 $\pm$ 0.2	40.3 $\pm$ 0.4	47.2 $\pm$ 0.6	10.4 $\pm$ 0.2	18.0 $\pm$ 0.2	19.7 $\pm$ 0.3	67.1 $\pm$ 1.3	77.0 $\pm$ 0.5	81.6 $\pm$ 0.9
NDPDC	26.6 $\pm$ 1.2	46.8 $\pm$ 0.6	53.9 $\pm$ 0.2	10.7 $\pm$ 0.2	17.5 $\pm$ 0.7	19.2 $\pm$ 0.3	67.0 $\pm$ 2.2	78.1 $\pm$ 1.1	82.3 $\pm$ 0.7
<b>DP-GENG (Ours)</b>	<b>35.6<math>\pm</math>0.9</b>	<b>59.0<math>\pm</math>0.3</b>	<b>65.5<math>\pm</math>0.4</b>	<b>19.3<math>\pm</math>0.3</b>	<b>27.9<math>\pm</math>0.4</b>	<b>32.3<math>\pm</math>0.2</b>	<b>69.3<math>\pm</math>0.4</b>	<b>81.4<math>\pm</math>0.4</b>	<b>85.7<math>\pm</math>0.3</b>
without DP guarantees									
DM	28.6 $\pm$ 0.6	48.9 $\pm$ 0.6	64.0 $\pm$ 0.4	11.4 $\pm$ 0.3	29.7 $\pm$ 0.3	43.6 $\pm$ 0.4	68.4 $\pm$ 0.6	80.1 $\pm$ 0.3	85.2 $\pm$ 0.4
NCFM	49.5 $\pm$ 0.3	71.8 $\pm$ 0.3	77.4 $\pm$ 0.3	34.4 $\pm$ 0.5	48.7 $\pm$ 0.3	54.7 $\pm$ 0.2	74.6 $\pm$ 0.4	84.8 $\pm$ 0.3	88.2 $\pm$ 0.4
Whole Dataset		84.8 $\pm$ 0.1			56.2 $\pm$ 0.3			95.6 $\pm$ 0.3	

Table 1: Comparison with previous methods on test accuracy (%). Results are averaged over three random seeds. IPC denotes images per class, and  $\epsilon$  is privacy budget. NCFM is the SOTA non-private distillation algorithm.

Method	CIFAR-10				CIFAR-100			
$\epsilon$	1		10		1		10	
Metric	Test Acc.	TPR@0.1%FPR	Test Acc.	TPR@0.1%FPR	Test Acc.	TPR@0.1%FPR	Test Acc.	TPR@0.1%FPR
PSG	35.9	0.08	47.2	0.13	10.3	0.09	19.7	0.20
NDPDC	42.6	0.10	53.9	0.14	11.5	0.11	19.2	0.18
<b>DP-GENG (Ours)</b>	55.9	0.10	65.5	0.14	25.9	0.12	32.3	0.17

Table 2: Comparison of different methods against MIA. For reference, a standard non-DP method (DM) achieves Test Acc./TPR@0.1%FPR of 63.0/0.82 on CIFAR-10 and 43.6/1.06 on CIFAR-100.

- DP Data Generator Methods: (i) Input level method: DP-MERF (Harder, Adamczewski, and Park 2021); (ii) Model-level method: PrivImage (Li et al. 2024a); (iii) Output level method: PE (Lin et al. 2023).

To illustrate the gap between DP-DD and standard DD algorithms, we select DM (Zhao and Bilen 2023) and NCFM (Wang et al. 2025) for comparison, neither of which has DP privacy guarantees.

**Experimental Setup.** Due to space constraints, we defer our implementation details to Appendix G (Shi et al. 2025).

## 4.2 Comparing Utility of Distilled Datasets

To comprehensively evaluate the effectiveness of our proposed method, we conduct experiments to compare the utility of our method against state-of-the-art approaches. We measure utility by the test accuracy of models trained on these distilled datasets. Table 1 presents the results across CIFAR-10, CIFAR-100, and CelebA datasets with varying privacy budgets and IPC settings. Higher test accuracy indicates

better utility preservation while maintaining DP guarantees.

Our method, DP-GENG, consistently achieves superior results under identical privacy budgets and IPC settings. For instance, on CIFAR-10 with  $\epsilon=10$  and IPC=50, DP-GENG attains 65.5% accuracy, markedly surpassing existing DP-DD approaches. Notably, directly generating DP data at the IPC scale yields lower utility than DP-DD methods, as it does not effectively distill information from private data. DP-GENG further bridges the performance gap to standard DD, significantly enhancing distilled data usability under DP guarantees. However, we note that on CIFAR-100, both DP-DD methods exhibit a significant performance gap compared to standard DD. We attribute this to the fact that CIFAR-100 has fewer samples per class. As a result, achieving DP guarantees requires injecting more noise, which reduces the utility of the distilled dataset.

To assess the generalizability of distilled datasets, we conducted extensive cross-architecture evaluations, with detailed results presented in Appendix H (Shi et al. 2025). DP-GENG demonstrates superior generalizability compared to other DP-

(a)			
DP-Init	DP-FM	DP-EG	Test Accuracy (%)
✓			48.7 $\pm$ 0.4
	✓		53.2 $\pm$ 0.2
✓	✓		60.8 $\pm$ 0.3
✓	✓	✓	<b>65.5<math>\pm</math>0.4</b>

(b)		
DP Generator	Test Accuracy (%)	
	CIFAR-10	CelebA
DP-MERF	56.4 $\pm$ 0.3	82.4 $\pm$ 0.2
PE	<b>65.5<math>\pm</math>0.4</b>	83.1 $\pm$ 0.3
PrivImage	64.7 $\pm$ 0.4	<b>85.7<math>\pm</math>0.2</b>

Table 3: Ablation studies. (a) Effectiveness of different components in DP-GENG (CIFAR-10, IPC=50,  $\epsilon = 10$ ). We progressively incorporate Initialization (Init), Feature Matching (FM), and Expert Guidance (EG) to demonstrate their individual contributions to the final performance. ✓ indicates the utilized component. (b) Comparison of different DP generators with IPC=50. We evaluate the impact of using different DP generators (DP-MERF, PE, and PrivImage) on the utility of the distilled dataset across CIFAR-10 and CelebA.

DD methods when evaluated on unseen architectures.

### 4.3 Comparing Privacy of Distilled Datasets Through Membership Inference Attacks

To evaluate privacy protection, we analyze our method’s resistance to MIAs by employing the LiRA (Carlini et al. 2022; Zhao and Zhang 2025b; Wang et al. 2024), strictly following their implementation. Table 2 illustrates the utility-privacy trade-off, where utility is measured by Test Accuracy and privacy by TPR@0.1%FPR. The results demonstrate that DP-GENG achieves comparable MIA resistance to other methods while offering superior data utility under the same privacy budget. This validates the correctness of DP-GENG’s overall privacy budget, derived via GDP composition (Lemma 2) and GDP-to-DP conversion (Lemma 3).

### 4.4 Ablation Study

**The Impact of Individual Components in DP-GENG.** To understand the contribution of each component in our method, we conduct an ablation study by selectively enabling different components. Table 3 presents the test accuracy on CIFAR-10 under different configurations. We analyze how each component (DP-Init, DP-FM, and DP-EG) contributes to the overall performance, demonstrating the necessity of our multi-stage approach for achieving optimal results.

**The Impact of DP Generators.** The DP-generated data required by DP-GENG can be produced by various DP generators. We investigate their impact on the final distilled dataset utility. Table 3 compares the test accuracies on CIFAR-10 and CelebA when DP-GENG utilizes DP-generated data from DP-MERF, PE, and PrivImage, all operating under the same privacy budget. The choice of DP generator can be dataset-dependent. For example, PE, as a training-free method, may exhibit reduced generation quality if there is a significant distributional divergence between the private and public data. This analysis helps identify suitable DP generators for maximizing utility while maintaining privacy guarantees.

**The Impact of Privacy Budget Allocation Strategy.** The allocation of privacy budgets across different components of DP-GENG is a crucial design choice. Table 4 presents an analysis of various privacy budget allocations on CIFAR-10 with IPC=50. The results indicate that our strategy can significantly improve the utility of the distilled dataset. Note that

$\epsilon_{\text{total}}$	$\mu_{\text{total}}^{\dagger}$	$\mu_G$	$\mu_F$	$\mu_E$	Test Accuracy (%)
10	2.00	1.21	1.17	1.17	57.2 $\pm$ 0.2
	2.00	0.27	1.07	1.66	59.7 $\pm$ 0.4
	2.00	0.27	1.75	0.92	61.3 $\pm$ 0.5
	<b>2.00</b>	<b>0.27</b>	<b>1.48</b>	<b>1.30</b>	<b>65.5<math>\pm</math>0.3</b>
20	<b>3.44</b>	<b>0.50</b>	<b>2.50</b>	<b>2.31</b>	<b>68.7<math>\pm</math>0.4</b>

Table 4: Ablation study on different privacy budget allocations. The bold rows are derived from our strategy. Results are for CIFAR-10 with IPC=50.  $\mu_{\text{total}}^{\dagger}$  is derived from the target  $\epsilon_{\text{total}}$  and  $\delta_{\text{total}}$ .

DP data generation can operate with a smaller privacy budget due to the capabilities of pre-trained generators. For training the expert model, allocating a large privacy budget would introduce excessive noise in the feature matching stage, while a small budget would lead to an inaccurate class distribution.

### 4.5 Visualization of Distilled Dataset

To provide qualitative insights into the performance of our method, we visualize the distilled datasets for CIFAR-10, CIFAR-100 and CelebA, generated by different DP-DD methods. In our extension version Appendix I (Shi et al. 2025), we present representative samples from PSG, NDPDC, and our method. We observe that DP-GENG generates data with much higher realism compared to the other methods, which enhances both data utility and cross-architecture generalization (Sun et al. 2024b; Shao et al. 2024).

## 5 Conclusion

This paper introduces DP-GENG, a novel framework for differentially private dataset distillation (DP-DD) that overcomes the utility and realism challenges faced by existing DP-DD methods. By strategically leveraging DP-generated data for initialization, training feature extractors, and employing an expert model for guided distillation, we observe that DP-GENG significantly outperforms state-of-the-art methods. Extensive experiments demonstrate DP-GENG’s superior performance in terms of dataset utility and robustness against privacy attacks, positioning it as a new, effective paradigm for trustworthy dataset distillation on sensitive data.

## Acknowledgments

This work was partly supported by the National Key Research and Development Program of China under No. 2024YFB3908400, NSFC under No. 62402418, 62402148, the Zhejiang Province's 2025 "Leading Goose + X" Science and Technology Plan under grant No.2025C02034, the Key R&D Program of Ningbo under No. 2024Z115, and the China Postdoctoral Science Foundation under No. 2024M762829.

## References

- Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H. B.; Mironov, I.; Talwar, K.; and Zhang, L. 2016. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 308–318.
- Carlini, N.; Chien, S.; Nasr, M.; Song, S.; Terzis, A.; and Tramer, F. 2022. Membership inference attacks from first principles. In *2022 IEEE Symposium on Security and Privacy (SP)*, 1897–1914. IEEE.
- Carlini, N.; Feldman, V.; and Nasr, M. 2022. No Free Lunch in "Privacy for Free: How does Dataset Condensation Help Privacy". *arXiv preprint arXiv:2209.14987*.
- Cazenavette, G.; Wang, T.; Torralba, A.; Efros, A. A.; and Zhu, J.-Y. 2022. Dataset distillation by matching training trajectories. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4750–4759.
- Chen, D.; Kerkouche, R.; and Fritz, M. 2022. Private set generation with discriminative information. *Advances in Neural Information Processing Systems*, 35: 14678–14690.
- Dong, J.; Roth, A.; and Su, W. J. 2022. Gaussian differential privacy. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 84(1): 3–37.
- Dong, T.; Zhao, B.; and Lyu, L. 2022. Privacy for free: How does dataset condensation help privacy? In *International Conference on Machine Learning*, 5378–5396. PMLR.
- Dwork, C.; McSherry, F.; Nissim, K.; and Smith, A. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*, 265–284. Springer.
- Dwork, C.; Roth, A.; et al. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4): 211–407.
- Farayola, O. A.; Olorunfemi, O. L.; and Shoetan, P. O. 2024. Data privacy and security in it: a review of techniques and challenges. *Computer Science & IT Research Journal*, 5(3): 606–615.
- Ghalebikesabi, S.; Berrada, L.; Goyal, S.; Ktena, I.; Stanforth, R.; Hayes, J.; De, S.; Smith, S. L.; Wiles, O.; and Balle, B. 2023. Differentially private diffusion models generate useful synthetic images. *arXiv preprint arXiv:2302.13861*.
- Guo, Z.; Wang, K.; Cazenavette, G.; Li, H.; Zhang, K.; and You, Y. 2024. Towards Lossless Dataset Distillation via Difficulty-Aligned Trajectory Matching. In *The Twelfth International Conference on Learning Representations*.
- Harder, F.; Adamczewski, K.; and Park, M. 2021. Dp-merf: Differentially private mean embeddings with random features for practical privacy-preserving data generation. In *International conference on artificial intelligence and statistics*, 1819–1827. PMLR.
- He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- Jayaraman, B.; and Evans, D. 2019. Evaluating differentially private machine learning in practice. In *28th USENIX Security Symposium (USENIX Security 19)*, 1895–1912.
- Jia, K.; Ma, Y.; Li, Y.; and Wang, F. 2025. PrAda-GAN: A Private Adaptive Generative Adversarial Network with Bayes Network Structure. *arXiv:2511.07997*.
- Karale, A. 2021. The Challenges of IoT Addressing Security, Ethics, Privacy, and Laws. *Internet Things*, 15: 100420.
- Kim, J.-H.; Kim, J.; Oh, S. J.; Yun, S.; Song, H.; Jeong, J.; Ha, J.-W.; and Song, H. O. 2022. Dataset condensation via efficient synthetic-data parameterization. In *International Conference on Machine Learning*, 11102–11118. PMLR.
- Krizhevsky, A.; Hinton, G.; et al. 2009. Learning multiple layers of features from tiny images.
- Lan, G.; Inan, H. A.; Abdelnabi, S.; Kulkarni, J.; Wutschitz, L.; Shokri, R.; Brinton, C. G.; and Sim, R. 2025. Contextual Integrity in LLMs via Reasoning and Reinforcement Learning. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS)*.
- LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11): 2278–2324.
- Li, G.; Togo, R.; Ogawa, T.; and Haseyama, M. 2020. Soft-label anonymous gastric x-ray image distillation. In *2020 IEEE International Conference on Image Processing (ICIP)*, 305–309. IEEE.
- Li, K.; Gong, C.; Li, Z.; Zhao, Y.; Hou, X.; and Wang, T. 2024a. {PrivImage}: Differentially Private Synthetic Image Generation using Diffusion Models with {Semantic-Aware} Pretraining. In *33rd USENIX Security Symposium (USENIX Security 24)*, 4837–4854.
- Li, Q.; Wang, C.-L.; Cao, Y.; and Wang, D. 2024b. Data Lineage Inference: Uncovering Privacy Vulnerabilities of Dataset Pruning. *arXiv preprint arXiv:2411.15796*.
- Lin, Z.; Gopi, S.; Kulkarni, J.; Nori, H.; and Yekhanin, S. 2023. Differentially private synthetic data via foundation model apis 1: Images. *arXiv preprint arXiv:2305.15560*.
- Liu, M. F.; Lyu, S.; Vinaroz, M.; and Park, M. 2023. Differentially private latent diffusion models. *arXiv preprint arXiv:2305.15759*.
- Liu, S.; Wang, K.; Yang, X.; Ye, J.; and Wang, X. 2022. Dataset Distillation via Factorization. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*.
- Liu, Z.; Luo, P.; Wang, X.; and Tang, X. 2015. Deep learning face attributes in the wild. In *Proceedings of the IEEE international conference on computer vision*, 3730–3738.

- Lu, W.; Tong, Y.; and Ye, Z. 2025. DAMMFND: Domain-Aware Multimodal Multi-view Fake News Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 559–567.
- Ma, Y.; and Yang, H. 2024. Optimal locally private nonparametric classification with public data. *Journal of Machine Learning Research*, 25(167): 1–62.
- Mironov, I. 2017. Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, 263–275. IEEE.
- Patwa, S.; Sun, D.; Gilad, A.; Machanavajjhala, A.; and Roy, S. 2023. DP-PQD: Privately Detecting Per-Query Gaps In Synthetic Data Generated By Black-Box Mechanisms. *arXiv preprint arXiv:2309.08574*.
- Regulation, P. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council. *Regulation (eu)*, 679: 2016.
- Sagun, L.; Evci, U.; Guney, V. U.; Dauphin, Y.; and Bottou, L. 2017. Empirical analysis of the hessian of over-parametrized neural networks. *arXiv preprint arXiv:1706.04454*.
- Shao, S.; Zhou, Z.; Chen, H.; and Shen, Z. 2024. Elucidating the design space of dataset condensation. *arXiv preprint arXiv:2404.13733*.
- Shi, S.; Zhang, J.; Jiang, S.; Zhou, C.; Li, Y.; Zhu, M.; Wu, Y.; and Du, T. 2025. DP-GENG : Differentially Private Dataset Distillation Guided by DP-Generated Data. *arXiv:2511.09876*.
- Shokri, R.; Stronati, M.; Song, C.; and Shmatikov, V. 2017. Membership inference attacks against machine learning models. In *2017 IEEE symposium on security and privacy (SP)*, 3–18. IEEE.
- Sun, D.; Chen, J. Q.; Gong, C.; Wang, T.; and Li, Z. 2024a. Netdpsyn: synthesizing network traces under differential privacy. In *Proceedings of the 2024 ACM on Internet Measurement Conference*, 545–554.
- Sun, P.; Shi, B.; Yu, D.; and Lin, T. 2024b. On the Diversity and Realism of Distilled Dataset: An Efficient Dataset Distillation Paradigm. In *CVPR*.
- Tong, Y.; Lu, W.; Cui, X.; Mao, Y.; and Zhao, Z. 2025. DAPT: Domain-Aware Prompt-Tuning for Multimodal Fake News Detection. In *Proceedings of the 33rd ACM International Conference on Multimedia*, 7902–7911.
- Tong, Y.; Lu, W.; Zhao, Z.; Lai, S.; and Shi, T. 2024. MDMFND: Multi-modal multi-domain fake news detection. In *Proceedings of the 32nd ACM International Conference on Multimedia*, 1178–1186.
- Wang, S.; Yang, Y.; Liu, Z.; Sun, C.; Hu, X.; He, C.; and Zhang, L. 2025. Dataset Distillation with Neural Characteristic Function: A Minmax Perspective. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Wang, T.; Zhu, J.-Y.; Torralba, A.; and Efros, A. A. 2018. Dataset distillation. *arXiv preprint arXiv:1811.10959*.
- Wang, Y.; Wang, X.; Lin, Z.; Su, Y.; Zhang, S.; Hou, R.; and Meng, D. 2024. Garrison: A high-performance gpu-accelerated inference system for adversarial ensemble defense. In *Proceedings of the 61st ACM/IEEE Design Automation Conference*, 1–6.
- Xiao, H.; Rasul, K.; and Vollgraf, R. 2017. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*.
- Yu, R.; Liu, S.; and Wang, X. 2024. Dataset distillation: A comprehensive review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Yuan, B.; Wang, Z.; Baktashmotlagh, M.; Luo, Y.; and Huang, Z. 2024. Color-Oriented Redundancy Reduction in Dataset Distillation. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*.
- Zhang, H.; Li, S.; Lin, F.; Wang, W.; Qian, Z.; and Ge, S. 2024. DANCE: Dual-View Distribution Alignment for Dataset Condensation. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.
- Zhang, L.; Zhang, J.; Lei, B.; Mukherjee, S.; Pan, X.; Zhao, B.; Ding, C.; Li, Y.; and Dongkuan, X. 2023. Accelerating Dataset Distillation via Model Augmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11950–11959.
- Zhao, B.; and Bilen, H. 2023. Dataset condensation with distribution matching. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 6514–6523.
- Zhao, B.; Mopuri, K. R.; and Bilen, H. 2020. Dataset condensation with gradient matching. *arXiv preprint arXiv:2006.05929*.
- Zhao, G.; Li, G.; Qin, Y.; and Yu, Y. 2023. Improved Distribution Matching for Dataset Condensation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7856–7865.
- Zhao, Y.; and Zhang, J. 2025a. Does Training with Synthetic Data Truly Protect Privacy? In *The Thirteenth International Conference on Learning Representations*.
- Zhao, Y.; and Zhang, J. 2025b. Does Training with Synthetic Data Truly Protect Privacy? In *The Thirteenth International Conference on Learning Representations*. GitHub repository.
- Zhao, Z. 2024. Balf: Simple and efficient blur aware local feature detector. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 3362–3372.
- Zhao, Z.; and Chen, B. M. 2023. Benchmark for Evaluating Initialization of Visual-Inertial Odometry. In *2023 42nd Chinese Control Conference (CCC)*, 3935–3940. IEEE.
- Zheng, T.; and Li, B. 2024. Differentially Private Dataset Condensation. In *Proceedings of the Network and Distributed System Security Symposium (NDSS), Workshop*.