Data-Free Knowledge Distillation of Deep Generative Models

Indian Institute of Technology Mandi

Group - 2

B23158, B23153, B23162, B23173, B23229

Mentor Ms. Kajal

Presented to:

Dr. Arnav Bhavsar
Dr. Prabhat Bharti

May 15, 2025

Outline

Introduction

Problem Definition

- Data-Free Knowledge Distillation (DFKD) enables transferring knowledge from a teacher network to a student network without access to original training data.
- Existing DFKD methods, including DiffDFKD[1], focus primarily on classification models.
- **Our project aims to generalize DFKD to any model with a continuous latent space**, such as VAEs, GANs, and diffusion models.
- We develop a more general DFKD pipeline that first synthesizes artificial dataset and then distills generative models.

[1] Qi, et al., Data-free Knowledge Distillation with Diffusion Models.arXiv:2504.00870v1 (April 2025)

Motivation

- **Why Data-Free Distillation?**
  - Data privacy, proprietary restrictions, and large-scale datasets make retaining or sharing data impractical.
  - DFKD allows efficient model compression and knowledge transfer without compromising data security.
- **Why extend to continuous latent space models?**
  - Current solutions (e.g., DiffDFKD) are task-specific, lacking generalizability to diverse generative models.
  - A unified DFKD framework would enable lightweight deployment of generative models on edge devices, privacy-critical environments, and low-resource platforms.

Our Objectives

- Analyze existing distillation strategies for generative models.
- Design a modular data free distillation pipeline applicable to any generative model with an encoder to a continuous latent space (VAE, Diffusion, GANs, etc.).
- Compare with-data and data-free distillation performance on standard datasets.

Related Works

Papers Referred for Methodology

Papers for Distillation for DGMs

- Distilling the Knowledge of a Neural Network[2]
- A Unified Knowledge Distillation Framework for Deep Directed Graphical Models (2023) [3]

Papers for Data Free Generation of Images

- Data-Free Knowledge Distillation for Deep Neural Networks[4]
- Data-free Knowledge Distillation with Diffusion Models (2025)[5]

---

[2] Hinton, Geoffrey Vinyals, Oriol Dean, Jeff. (2015). Distilling the Knowledge in a Neural Network.

[3] Y. Chen, K. Liang, Z. Zeng, S. Yao and H. Shao, "A Unified Knowledge Distillation Framework for Deep Directed Graphical Models," 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, pp. 7795-7804, doi: 10.1109/CVPR52729.2023.00753.

[4] Lopes, Raphael Fenu, Stefano Starner, Thad. (2017). Data-Free Knowledge Distillation for Deep Neural Networks. 10.48550/arXiv.1710.07535.

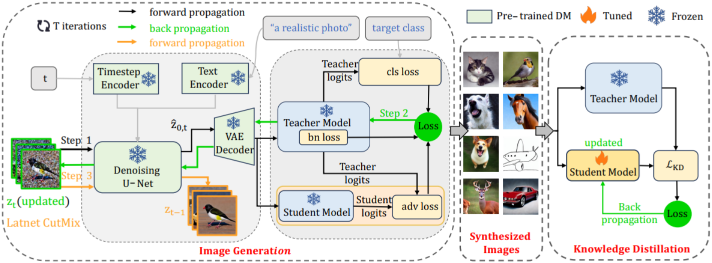
[5] Qi, et al., Data-free Knowledge Distillation with Diffusion Models.arXiv:2504.00870v1 (April 2025)

Related Works

Method	Datafree Generation	Distillation	DGM Generalization
Hinton, Geoffrey	No	Yes	No
Lopes, Raphael	Yes	Yes	No
Y. Chen, K. Liang	No	Yes	Yes
Qi, et al.	Yes	Yes	No

Methodology

Data Free Generation



[6] Data Free Generation of Images through inferencing the Teacher Model and further Distillation(done on a Classifier)
The Losses:-

- Batch Normalization-Matching image logits with teacher distribution.
- Classification Loss-The teacher should classify images from its dataset correctly.
- Adveresarial Loss- Pushes student and teacher model away for generalization.

[6] Taken from DiffDFKD(Qi et. al , april 2025)

Our Method

We mainly focus on VAEs for our experimentation and the method is as follows:-

Data Free Generation through Model Inference

Creation of dataset by extracting feature from our pretrained VAE

- Through conditional prompts we guide our model to generate images of a certain image space.
- Through our losses we update the input vectors and force the images generated to match our teacher model's distribution.
- Losses being the KL Divergence loss and the MSE loss between the generated image by the diffusion model and the reconstruction.

Distillation of DGM's

[7] For our demonstration we focus on the Distillation of VAEs using the Unified Distillation Model Framework for VAEs.

[7] Taken From Y. Chen, K. Liang, Z. Zeng, S. Yao and H. Shao, "A Unified Knowledge Distillation Framework for Deep Directed Graphical Models," 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, pp. 7795-7804, doi: 10.1109/CVPR52729.2023.00753.

The Pipeline Created

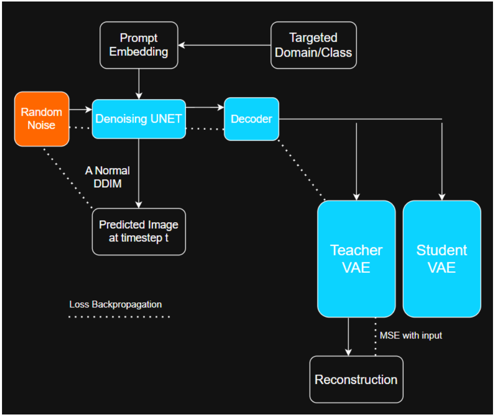


Figure: Pipeline for data synthesis. After data synthesis, we follow the DGM distillation process as described in Chen *et al.*, 2023.

Experiments

Setup

- Datasets Used
  - CIFAR-10: 60,000 images (32x32x3) belonging to 10 classes
  - CelebA-HQ: 30,000 images of human faces (1024x1024x3)
- Models Used
  - Stable-Diffusion (v1.5) for guided synthesis
  - LVAE: Heirachial VAE ($\sim$5.6M Params) as pre-trained teacher
  - Student LVAE, similar architecture but lesser parameters ($\sim$0.03M)
- 1000 unique images synthesized
- Distilled for 50 epochs with batch size 128

DGM distillation with CIFAR-10

Table: Results of Student VAE distilled using CIFAR-10

Model	#param	FID	EMD	MMD	1NN	Inference Time	Model Size (MB)
Teacher	5.39M	4.63	7.57	0.25	0.89	23.3±1.12ms	20.5
Student	0.03M	5.24	7.96	0.30	0.91	5.63±0.54ms	1.44

Data-synthesis (paper's)

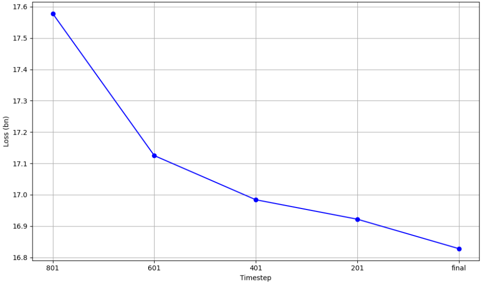


Figure: Batch Norm loss v/s timesteps

Figure: Some generations[a]

[a] Taken from DiffDFKD(Qi et. al, Apr 2025)

Data-synthesis from CelebA-HQ (ours)

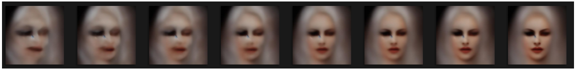
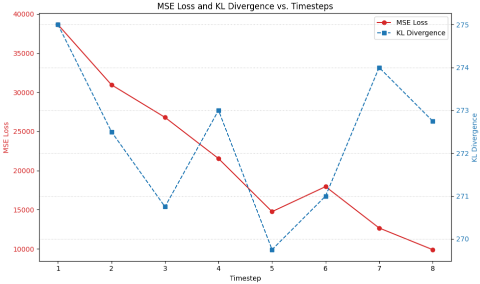
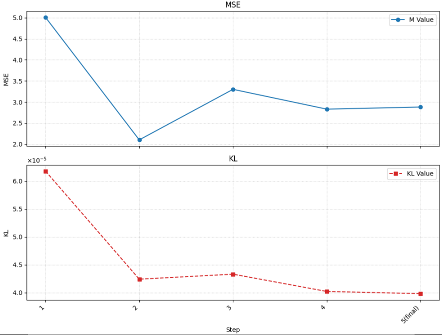


Figure: Example of CelebA synthesis using our pipeline



Data-synthesis from CIFAR-10 (ours)

Figure: Example of CIFAR-10 synthesis using our pipeline



Qualitative Results from our pipeline

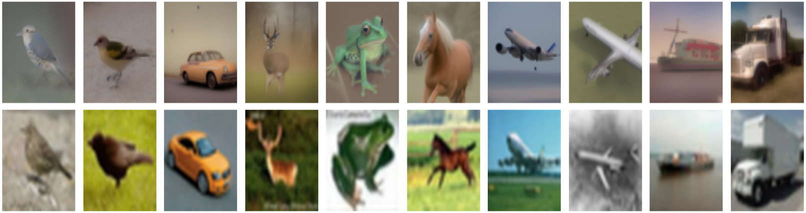


Figure: Synthesized images from our pipeline (1st row) v/s their closest match in CIFAR-10 dataset (2nd row)

Data-free DGM distillation (our pipeline)

Table: Results of DGM distillation via our data-free pipeline

Model	#param	FID	EMD	MMD	1NN	Inference Time	Model Size (MB)
Teacher	5.39M	4.63	7.57	0.25	0.89	23.3±1.12ms	20.5
Student (with CIFAR-10)	0.03M	5.24	7.96	0.30	0.91	5.63±0.54ms	1.44
Student (data-free)	0.03M	5.27	8.07	0.29	0.92	5.6±1.2ms	1.52

Conclusion

Conclusion

- Implemented Knowledge Distillation pipeline for DGMs and tested on LVAE.
- Implemented Data-free KD pipeline using Stable Diffusion for Classification tasks.
- Developed a pipeline for Data-free KD for DGMs using Stable Diffusion for Quality data-synthesis.

Future work and Limitations

1. The pipeline can be tested on other generative DGMs like Diffusion models.
2. We can't exactly get our losses to be very small due to approximate inference used in many Generative Models.

Thank You!