Novelty Assessment Report

Paper: Provable Separations between Memorization and Generalization in Diffusion Models

PDF URL: https://openreview.net/pdf?id=42gfTZzyvV

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-07

Abstract

Diffusion models have achieved remarkable success across diverse domains, but they remain vulnerable to memorization---reproducing training data rather than generating novel outputs. This not only limits their creative potential but also raises concerns about privacy and safety. While empirical studies have explored mitigation strategies, theoretical understanding of memorization remains limited. We address this gap through developing a dual-separation result via two complementary perspectives: statistical estimation and network approximation. From the estimation side, we show that the ground-truth score function does not minimize the empirical denoising loss, creating a separation that drives memorization. From the approximation side, we prove that implementing the empirical score function requires network size to scale with sample size, spelling a separation compared to the more compact network representation of the ground-truth score function. Guided by these insights, we develop a pruning-based method that reduces memorization while maintaining generation quality in diffusion transformers.

Core Task Landscape

This paper addresses: **memorization versus generalization in diffusion models**

A total of **50 papers** were analyzed and organized into a taxonomy with **19 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Theoretical Foundations and Mechanisms**
- **Empirical Characterization and Detection**
- **Mitigation and Prevention Strategies**
- **Memorization-Generalization Trade-offs and Relationships**
- **Application-Specific Memorization and Generalization**

Complete Taxonomy Tree

- memorization versus generalization in diffusion models Survey Taxonomy
- Theoretical Foundations and Mechanisms
  - Score Matching and Empirical Loss Analysis (3 papers)
  - [1] From memorization to generalization: a theoretical framework for diffusion-based generative models (Halder, 2024) View paper
  - [2] Memorization and Regularization in Generative Diffusion Models (Baptista, 2025) View paper
  - [13] On the Generalization of Diffusion Model (Yi, 2023) View paper
  - Manifold and Latent Structure Theory (3 papers)
  - [3] Memorization and generalization in generative diffusion under the manifold hypothesis (Beatrice Achilli, 2025) View paper
  - [21] Generalization in diffusion models arises from geometry-adaptive harmonic representations (Kadkhodaie, 2023) View paper
  - [24] Understanding generalizability of diffusion models requires rethinking the hidden gaussian structure (Yixiang Dai, 2024) View paper
  - Network Capacity and Approximation Separations ★ (3 papers)
  - [0] Provable Separations between Memorization and Generalization in Diffusion Models (Anon et al., 2026) View paper
  - [29] Generalization of diffusion models: Principles, theory, and implications (H Zhang, 2025) View paper
  - [49] Overparameterization and double descent in PCA, GANs, and Diffusion models (Luzi, 2024) View paper
  - Associative Memory and Dynamical Systems Perspectives (2 papers)
  - [4] Memorization to Generalization: Emergence of Diffusion Models from Associative Memory (Pham Bao, 2025) View paper
  - [35] Generative diffusion models from a PDE perspective (Cao Fei, 2025) View paper
- Empirical Characterization and Detection
  - Memorization Detection and Measurement Methods (3 papers)
  - [8] Detecting, Explaining, and Mitigating Memorization in Diffusion Models (Wen, 2024) View paper
  - [9] On the discrepancy and connection between memorization and generation in diffusion models (H Wang, 2024) View paper
  - [46] MemBench: Memorized Image Trigger Prompt Dataset for Diffusion Models (Oh, 2024) View paper
  - Localization and Spatial Analysis of Memorization (4 papers)
  - [10] Memorization is Localized within a Small Subspace in Diffusion Models (R Chavhan, 2024) View paper
  - [15] Finding NeMo: Localizing Neurons Responsible For Memorization in Diffusion Models (Franziska Boenisch, 2024) View paper
  - [19] Exploring local memorization in diffusion models via bright ending attention (Chen Chen, 2024) View paper
  - [32] Demystifying Foreground-Background Memorization in Diffusion Models (Di, 2025) View paper
  - Cross-Attention and Prompt-Conditioned Memorization (2 papers)
  - [11] Tracking memorization geometry throughout the diffusion model generative process (J Brokman, 2025) View paper
  - [20] Unveiling and Mitigating Memorization in Text-to-image Diffusion Models through Cross Attention (Ren Jie, 2024) View paper

    - Empirical Observations of Memorization-Generalization Transitions (5 papers)
    - [7] On Memorization in Diffusion Models (Gu Xiangming, 2023) View paper
    - [14] Memorization vs. Generalization in diffusion models with the U-Net architecture (Topeza, 2025) View paper
    - [16] Early-stopping Too Late? Traces of Memorization Before Overfitting in Generative Diffusion (J Garnier-Brun, 2025) View paper
    - [18] Bigger Isn't Always Memorizing: Early Stopping Overparameterized Diffusion Models (Favero, 2025) View paper
    - [26] Diffusion probabilistic models generalize when they fail to memorize (TH Yoon, 2023) View paper
- Mitigation and Prevention Strategies
    - Training-Time Regularization and Data Interventions (4 papers)
    - [12] Towards Memorization-Free Diffusion Models (Chen Chen, 2024) View paper
    - [23] Does generation require memorization? creative diffusion models using ambient diffusion (Shah, 2025) View paper
    - [39] Redistribute Ensemble Training for Mitigating Memorization in Diffusion Models (Xiaoliu Guan, 2025) View paper
    - [48] Data Cartography for Detecting Memorization Hotspots and Guiding Data Interventions in Generative Models (Laksh Patel, 2025) View paper
    - Parameter-Efficient Fine-Tuning and Capacity Control (2 papers)
    - [31] MemControl: Mitigating Memorization in Diffusion Models via Automated Parameter Selection (Raman Dutt, 2025) View paper
    - [34] Why Diffusion Models Don't Memorize: The Role of Implicit Dynamical Regularization in Training (Bonnaire, 2025) View paper
    - Inference-Time Guidance and Privacy-Preserving Generation (2 papers)
    - [33] Enhancing Privacy-Utility Trade-offs to Mitigate Memorization in Diffusion Models (Chen Chen, 2025) View paper
    - [36] Beyond Memorization: Gradient Projection Enables Selective Learning in Diffusion Models (Divya Kothandaraman, 2025) View paper
- Memorization-Generalization Trade-offs and Relationships
    - Rethinking Trade-offs and Mutual Relationships (3 papers)
    - [5] Rethinking Memorizationâ��Generalization Trade-Off in Generative Models (J Chae, 2025) View paper
    - [6] On the Edge of Memorization in Diffusion Models (Buchanan, 2025) View paper
    - [42] Understanding and Improving Diffusion Models (ìµ�ì£¼ì��, 2025) View paper
    - Model Collapse and Recursive Training Dynamics (1 papers)
    - [40] A Closer Look at Model Collapse: From a Generalization-to-Memorization Perspective (Shi Lianghe, 2025) View paper
    - Reproducibility and Cross-Model Consistency (2 papers)
    - [22] The emergence of reproducibility and generalizability in diffusion models (ZHANG Huijie, 2023) View paper
    - [30] Galaxy Imaging with Generative Models: Insights from a Two-Models Framework (Campagne, 2025) View paper
- Application-Specific Memorization and Generalization
    - Personalization and Subject-Specific Customization (3 papers)
    - [37] Infusion: Preventing Customized Text-to-Image Diffusion from Overfitting (Weili Zeng, 2024) View paper
    - [38] T-lora: Single image diffusion model customization without overfitting (Soboleva, 2025) View paper
    - [41] SSIE-Diffusion: Personalized Generative Model for Subject-Specific Image Editing (Guo Lin, 2024) View paper
    - Domain Adaptation and Transfer Learning (2 papers)
    - [17] Boosting domain generalized and adaptive detection with diffusion models: Fitness, generalization, and transferability (HE Boyong, 2025) View paper
    - [25] Unpaired Deblurring via Decoupled Diffusion Model (CHENG Junhao, 2025) View paper
    - Medical Imaging and Privacy-Sensitive Domains (1 papers)
    - [27] Investigating Data Memorization in 3D Latent Diffusion Models for Medical Image Synthesis (Salman Ul Hassan Dar, 2023) View paper
    - Discriminative Tasks and Non-Generative Applications (3 papers)
    - [43] DepthMaster: Taming Diffusion Models for Monocular Depth Estimation (Song, 2025) View paper
    - [44] Video Summarization using Denoising Diffusion Probabilistic Model (Zirui Shang, 2024) View paper
    - [45] Diffusion Model-Augmented Behavioral Cloning (Chen, 2023) View paper
    - Resource-Constrained and Deployment Considerations (3 papers)
    - [28] Prioritized Generative Replay (Wang, 2024) View paper
    - [47] Generative Adversarial Diffusion (U Jun, 2025) View paper
    - [50] Techniques for Maintaining Stability and High-Fidelity Outputs in Resource-Constrained Deployments of Large Generative Models (L Petersen, 2025) View paper

Narrative

Core task: memorization versus generalization in diffusion models. The field examines when and why diffusion models reproduce training data versus generating novel samples, organizing research into five main branches. Theoretical Foundations and Mechanisms investigates the underlying principles—such as network capacity limits, manifold hypotheses (Manifold Hypothesis Diffusion[3]), and emergent associative memory (Associative Memory Emergence[4])—that govern the memorization-generalization boundary. Empirical Characterization and Detection develops methods to identify and measure memorization, including geometric tracking (Tracking Memorization Geometry[11]) and localized subspace analysis (Localized Memorization Subspace[10]). Mitigation and Prevention Strategies proposes interventions like regularization techniques (Memorization and Regularization[2]), early stopping (Early-stopping Memorization Traces[16]), and memory-free training (Memorization-Free Diffusion[12]). The remaining branches explore trade-offs between privacy and utility (Privacy-Utility Trade-offs[33]) and application-specific challenges in domains such as medical imaging (3D Medical Memorization[27]) and video synthesis.

Recent work reveals contrasting perspectives on whether memorization is harmful or beneficial, with some studies framing it as a necessary phase (Memorization to Generalization Framework[1]) and others seeking to eliminate it entirely. A particularly active line examines capacity-driven separations: Memorization Generalization Separations[0] focuses on network approximation limits that create sharp boundaries between memorizing and generalizing regimes, closely aligning with theoretical investigations of overparameterization (Overparameterization Double Descent[49]) and capacity principles (Generalization Principles Theory[29]). Compared to empirical detection methods like Edge of Memorization[6], which characterize when models transition between regimes, Memorization Generalization Separations[0] emphasizes formal separations that reveal fundamental architectural constraints. This theoretical lens complements practical mitigation efforts, highlighting open questions about whether optimal generalization requires carefully tuned capacity or whether alternative training dynamics can bypass these trade-offs.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. Generalization of diffusion models: Principles, theory, and implications

Authors: H Zhang, P Wang, S Chen, Z Zhang, Q Qu | **Year/Venue**: 2025 | **URL**: View paper

Abstract

â¦ capacity is insufficient for memorization, diffusion models generate â¦ in generative models and diffusion models, including the generalizability and low-dimensional structures in diffusion â¦

Relationship Analysis

Both papers belong to the Network Capacity and Approximation Separations category, examining theoretical foundations of memorization versus generalization through network architecture requirements. They overlap in analyzing how network capacity influences the memorization-generalization trade-off in diffusion models. The original paper provides dual-separation results through statistical estimation and neural approximation perspectives with specific bounds on network size scaling with sample size, while the candidate paper appears to focus more broadly on generalization principles and low-dimensional structures in diffusion models.

---

2. Overparameterization and double descent in PCA, GANs, and Diffusion models

Authors: L Luzi | **Year/Venue**: 2024 | **URL**: View paper

Abstract

â¦ generative models that learn distributions by minimizing a metric or f-divergence do not exhibit double descent in generalization â¦ overfitting similar to what we saw in the linear models. â¦

Relationship Analysis

Both papers belong to the Network Capacity and Approximation Separations category, examining theoretical aspects of network size requirements in diffusion models. The original paper provides a dual-separation framework proving that ground-truth score functions require compact networks while empirical score functions need networks scaling with sample size, and connects this to memorization versus generalization. The candidate paper (a PhD thesis) focuses on overparameterization and double descent phenomena across multiple generative models including diffusion models, investigating how model capacity affects generalization through pseudo-supervision methods rather than establishing formal separation bounds between memorization and generalization regimes.

Contributions Analysis

Overall novelty summary. The paper develops a dual-separation theory for memorization in diffusion models, combining statistical estimation and network approximation perspectives. It resides in the 'Network Capacity and Approximation Separations' leaf under 'Theoretical Foundations and Mechanisms', which contains only three papers total. This represents a relatively sparse research direction within the broader taxonomy of 50 papers across 36 topics, suggesting the theoretical analysis of capacity-driven memorization boundaries remains an emerging area compared to more crowded empirical detection branches.

The taxonomy tree reveals neighboring theoretical branches examining manifold hypotheses and associative memory perspectives, alongside empirical characterization methods like geometric tracking and localized subspace analysis. The paper's dual-perspective approach bridges these areas: its statistical separation connects to score matching theory, while its approximation results relate to capacity principles explored in sibling work on overparameterization. The scope notes clarify that this leaf focuses specifically on formal separations via network size requirements, distinguishing it from empirical overfitting studies or training dynamics analyses found elsewhere.

Among 19 candidates examined across three contributions, no clearly refuting prior work was identified. The statistical separation theory examined 10 candidates with zero refutable matches, while the neural architectural separation theory examined 9 candidates, also with zero refutations. The pruning-based mitigation method was not evaluated against prior work in this limited search. These statistics suggest that within the top-K semantic matches and citation expansion performed, the dual-separation framework appears distinct from existing theoretical characterizations, though the search scope remains constrained.

Based on the limited literature search covering 19 candidates, the dual-perspective theoretical framework appears to occupy a relatively unexplored position within capacity-driven memorization analysis. The sparse population of its taxonomy leaf and absence of refuting candidates among examined papers suggest novelty, though a more exhaustive search across the broader theoretical foundations branch would strengthen this assessment.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Statistical separation theory for memorization in diffusion models

Description: The authors establish that the ground-truth score function does not minimize the empirical denoising score matching loss, creating a statistical gap that drives memorization. For mixture models, they provide a lower bound on this gap, formally characterizing how memorization arises from a statistical perspective.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Regularizing score-based models with score fokker-planck equations

URL: View paper

Brief Assessment

Regularizing Score Fokker-Planck[57] focuses on regularizing score-based models using Fokker-Planck equations to enforce physical consistency, not on analyzing statistical gaps in denoising score matching or memorization phenomena in diffusion models.

---

2. Optimizing Input of Denoising Score Matching is Biased Towards Higher Score Norm

URL: View paper

Brief Assessment

Optimizing Input Bias[58] focuses on bias in denoising score matching when optimizing conditional inputs or data distributions, not on the statistical gap between ground-truth and empirical score functions that drives memorization in finite-sample regimes.

---

3. Denoising Likelihood Score Matching for Conditional Score-based Data Generation

URL: View paper

Brief Assessment

Denoising Likelihood Matching[56] focuses on conditional score-based generation using classifier guidance, not on the statistical gap between ground-truth and empirical scores in denoising score matching that drives memorization.

---

4. Predicting molecular conformation via dynamic graph score matching

URL: View paper

Brief Assessment

Dynamic Graph Matching[52] focuses on molecular conformation prediction using score matching for spatial coordinates, not on analyzing statistical gaps or memorization in diffusion models for general data distributions.

5. Beyond point prediction: Score matching-based pseudolikelihood estimation of neural marked spatio-temporal point process

URL: View paper

Brief Assessment

Neural Marked Spatio-temporal[55] focuses on spatio-temporal point processes using score matching for pseudolikelihood estimation, not on diffusion models' memorization or statistical gaps in denoising score matching.

6. Taking a Big Step: Large Learning Rates in Denoising Score Matching Prevent Memorization

URL: View paper

Brief Assessment

Large Learning Rates[59] focuses on implicit regularization via learning rates in denoising score matching, not on establishing statistical gaps between ground-truth and empirical scores as the original paper does.

7. Fp-diffusion: Improving score-based diffusion models by enforcing the underlying score fokker-planck equation

URL: View paper

Brief Assessment

Fp-diffusion[54] focuses on enforcing the score Fokker-Planck equation to improve diffusion models, not on analyzing statistical gaps in denoising score matching or memorization phenomena. The candidate addresses model consistency and likelihood improvement through PDE-based regularization, which is a different technical approach from studying memorization via statistical estimation theory.

8. Evaluating the design space of diffusion-based generative models

URL: View paper

Brief Assessment

Evaluating Design Space[53] focuses on training convergence and sampling error analysis for diffusion models, not on the statistical gap between ground-truth and empirical score functions that drives memorization.

9. Maximum Likelihood Training for Score-Based Diffusion ODEs by High-Order Denoising Score Matching

URL: View paper

Brief Assessment

Maximum Likelihood Training[51] focuses on the gap between score matching objectives and maximum likelihood for score-based diffusion ODEs, not on the statistical gap between ground-truth and empirical score functions that drives memorization in finite-sample regimes.

10. Test-time Conditional Text-to-Image Synthesis Using Diffusion Models

URL: View paper

Brief Assessment

Test-time Conditional Synthesis[60] focuses on conditional image generation using test-time manipulation of noise predictions, not on statistical analysis of memorization or score matching loss gaps in diffusion models.

Contribution 2: Neural architectural separation theory for score function representation

Description: The authors prove that the ground-truth score function admits a compact neural representation, whereas approximating the empirical score function requires network size to scale with the sample size. This reveals a fundamental separation in the approximation capacity needed for these two functions.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Approximation and estimation bounds for artificial neural networks

URL: View paper

Brief Assessment

Artificial Neural Networks[70] focuses on approximation bounds for general function classes using neural networks, not specifically on score functions in diffusion models or the separation between ground-truth versus empirical score functions.

2. Approximation error of Fourier neural networks

URL: View paper

Brief Assessment

Fourier Neural Networks[69] focuses on approximation error bounds for Fourier neural networks with cosine activation functions, not on score function representation in diffusion models or the separation between ground-truth and empirical score functions.

3. Generalization error bound for denoising score matching under relaxed manifold assumption

URL: View paper

Brief Assessment

Relaxed Manifold Assumption[64] focuses on generalization error bounds for denoising score matching under a relaxed manifold assumption with Gaussian mixtures, not on neural network approximation separations between ground-truth and empirical score functions.

4. Generalization bounds for score-based generative models: a synthetic proof

URL: View paper

Brief Assessment

Synthetic Proof Bounds[68] focuses on generalization bounds for score-based generative models using neural network approximation theory, but does not establish a separation between network sizes needed for ground-truth versus empirical score functions as claimed in the original contribution.

5. Score approximation, estimation and distribution recovery of diffusion models on low-dimensional data

URL: View paper

Brief Assessment

Low-dimensional Data Recovery[66] focuses on score approximation for low-dimensional linear subspace data using encoder-decoder architectures, not on comparing network size requirements between ground-truth and empirical score functions as the original paper does.

---

6. Optimal estimation of a factorizable density using diffusion models with ReLU neural networks

URL: View paper

Brief Assessment

Factorizable Density Estimation[62] focuses on approximating score functions for factorizable densities with low-dimensional structure, not on comparing network size requirements between ground-truth and empirical score functions as in the original paper.

---

7. On the approximation of functions by tanh neural networks

URL: View paper

Brief Assessment

Tanh Neural Networks[61] focuses on approximation bounds for general Sobolev-regular and analytic functions using tanh activation, not specifically on score functions in diffusion models or the separation between ground-truth versus empirical score representations.

---

8. Approximation of RKHS Functionals by Neural Networks

URL: View paper

Brief Assessment

RKHS Functionals Approximation[63] focuses on approximating functionals on reproducing kernel Hilbert spaces using neural networks, not on score functions in diffusion models. The technical domains are distinct.

---

9. Convergence analysis of probability flow ode for score-based generative models

URL: View paper

Brief Assessment

Probability Flow Convergence[67] focuses on convergence analysis of probability flow ODEs for score-based generative models, not on neural network approximation bounds for ground-truth versus empirical score functions.

Contribution 3: Pruning-based mitigation method for diffusion transformers

Description: Guided by their theoretical insights, the authors propose a practical pruning method that identifies and removes attention heads contributing least in the small-time regime. This approach reduces memorization while preserving generation quality, validated through experiments on CIFAR-10.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Provable Separations between Memorization and Generalization in Diffusion Models View paper
- [1] From memorization to generalization: a theoretical framework for diffusion-based generative models View paper
- [2] Memorization and Regularization in Generative Diffusion Models View paper
- [3] Memorization and generalization in generative diffusion under the manifold hypothesis View paper
- [4] Memorization to Generalization: Emergence of Diffusion Models from Associative Memory View paper
- [5] Rethinking Memorizationâ Generalization Trade-Off in Generative Models View paper
- [6] On the Edge of Memorization in Diffusion Models View paper
- [7] On Memorization in Diffusion Models View paper
- [8] Detecting, Explaining, and Mitigating Memorization in Diffusion Models View paper
- [9] On the discrepancy and connection between memorization and generation in diffusion models View paper
- [10] Memorization is Localized within a Small Subspace in Diffusion Models View paper
- [11] Tracking memorization geometry throughout the diffusion model generative process View paper
- [12] Towards Memorization-Free Diffusion Models View paper
- [13] On the Generalization of Diffusion Model View paper
- [14] Memorization vs. Generalization in diffusion models with the U-Net architecture View paper
- [15] Finding NeMo: Localizing Neurons Responsible For Memorization in Diffusion Models View paper
- [16] Early-stopping Too Late? Traces of Memorization Before Overfitting in Generative Diffusion View paper
- [17] Boosting domain generalized and adaptive detection with diffusion models: Fitness, generalization, and transferability View paper
- [18] Bigger Isn't Always Memorizing: Early Stopping Overparameterized Diffusion Models View paper
- [19] Exploring local memorization in diffusion models via bright ending attention View paper
- [20] Unveiling and Mitigating Memorization in Text-to-image Diffusion Models through Cross Attention View paper
- [21] Generalization in diffusion models arises from geometry-adaptive harmonic representations View paper
- [22] The emergence of reproducibility and generalizability in diffusion models View paper
- [23] Does generation require memorization? creative diffusion models using ambient diffusion View paper
- [24] Understanding generalizability of diffusion models requires rethinking the hidden gaussian structure View paper
- [25] Unpaired Deblurring via Decoupled Diffusion Model View paper
- [26] Diffusion probabilistic models generalize when they fail to memorize View paper
- [27] Investigating Data Memorization in 3D Latent Diffusion Models for Medical Image Synthesis View paper
- [28] Prioritized Generative Replay View paper
- [29] Generalization of diffusion models: Principles, theory, and implications View paper
- [30] Galaxy Imaging with Generative Models: Insights from a Two-Models Framework View paper
- [31] MemControl: Mitigating Memorization in Diffusion Models via Automated Parameter Selection View paper

- [32] Demystifying Foreground-Background Memorization in Diffusion Models View paper
- [33] Enhancing Privacy-Utility Trade-offs to Mitigate Memorization in Diffusion Models View paper
- [34] Why Diffusion Models Don't Memorize: The Role of Implicit Dynamical Regularization in Training View paper
- [35] Generative diffusion models from a PDE perspective View paper
- [36] Beyond Memorization: Gradient Projection Enables Selective Learning in Diffusion Models View paper
- [37] Infusion: Preventing Customized Text-to-Image Diffusion from Overfitting View paper
- [38] T-lora: Single image diffusion model customization without overfitting View paper
- [39] Redistribute Ensemble Training for Mitigating Memorization in Diffusion Models View paper
- [40] A Closer Look at Model Collapse: From a Generalization-to-Memorization Perspective View paper
- [41] SSIE-Diffusion: Personalized Generative Model for Subject-Specific Image Editing View paper
- [42] Understanding and Improving Diffusion Models View paper
- [43] DepthMaster: Taming Diffusion Models for Monocular Depth Estimation View paper
- [44] Video Summarization using Denoising Diffusion Probabilistic Model View paper
- [45] Diffusion Model-Augmented Behavioral Cloning View paper
- [46] MemBench: Memorized Image Trigger Prompt Dataset for Diffusion Models View paper
- [47] Generative Adversarial Diffusion View paper
- [48] Data Cartography for Detecting Memorization Hotspots and Guiding Data Interventions in Generative Models View paper
- [49] Overparameterization and double descent in PCA, GANs, and Diffusion models View paper
- [50] Techniques for Maintaining Stability and High-Fidelity Outputs in Resource-Constrained Deployments of Large Generative Models View paper
- [51] Maximum Likelihood Training for Score-Based Diffusion ODEs by High-Order Denoising Score Matching View paper
- [52] Predicting molecular conformation via dynamic graph score matching View paper
- [53] Evaluating the design space of diffusion-based generative models View paper
- [54] Fp-diffusion: Improving score-based diffusion models by enforcing the underlying score fokker-planck equation View paper
- [55] Beyond point prediction: Score matching-based pseudolikelihood estimation of neural marked spatio-temporal point process View paper
- [56] Denoising Likelihood Score Matching for Conditional Score-based Data Generation View paper
- [57] Regularizing score-based models with score fokker-planck equations View paper
- [58] Optimizing Input of Denoising Score Matching is Biased Towards Higher Score Norm View paper
- [59] Taking a Big Step: Large Learning Rates in Denoising Score Matching Prevent Memorization View paper
- [60] Test-time Conditional Text-to-Image Synthesis Using Diffusion Models View paper
- [61] On the approximation of functions by tanh neural networks View paper
- [62] Optimal estimation of a factorizable density using diffusion models with ReLU neural networks View paper
- [63] Approximation of RKHS Functionals by Neural Networks View paper
- [64] Generalization error bound for denoising score matching under relaxed manifold assumption View paper
- [65] Approximation of compositional functions with ReLU neural networks View paper
- [66] Score approximation, estimation and distribution recovery of diffusion models on low-dimensional data View paper
- [67] Convergence analysis of probability flow ode for score-based generative models View paper
- [68] Generalization bounds for score-based generative models: a synthetic proof View paper
- [69] Approximation error of Fourier neural networks View paper
- [70] Approximation and estimation bounds for artificial neural networks View paper