DTU Compute
Department of Applied Mathematics and Computer Science

PIONEER CENTRE FOR
ARTIFICIAL INTELLIGENCE

Latent Diffusion for Missing Data

Alberte Heering Estad[1] **Ignacio Peis**[1,2] **Jes Frellsen**[1,2]

[1] Technical University of Denmark · [2] Pioneer Centre for Artificial Intelligence

1 Introduction

Real-world data are often incomplete, yet most diffusion-based imputers operate in *pixel space* and assume complete training data. They degrade sharply when it is not available.

Does moving diffusion into a learned latent space improve robustness under missing-completely-at-random (MCAR) corruption? We introduce **LDMiss** and benchmark it against pixel-space diffusion on MNIST across training missing rates.

Key Contributions

- First systematic comparison of **pixel- versus latent-space** diffusion trained on incomplete data.
- **LDMiss,** a robust β-VAE imputer paired with a latent-space diffusion prior.
- A **self-guidance** imputation scheme for unconditional latent diffusion.
- Best imputation across *all* training missing rates, robust even at 80% missing.

2 Score-Based Diffusion

Score-based models corrupt data toward noise through a forward SDE and generate by simulating the **reverse SDE**, which depends only on the learned score function:

$$d\mathbf{x} = \left[\,\mathbf{f}(\mathbf{x},t) - g(t)^2\nabla_{\mathbf{x}}\log p_t(\mathbf{x})\,\right]dt + g(t)\,d\bar{\mathbf{w}}$$

We use the Variance-Preserving SDE and train the score network by **denoising score matching**. LDMiss simply relocates this standard machinery to latent space.

3 Latent Diffusion (LDMiss)

A **β-VAE** with a KL weight of 10^{-6} compresses each 28×28 image to a **7×7×2 latent**, a reduction from 784 to 98 values, or **8× smaller**. Diffusion runs entirely in this space, so the score network never processes raw zero-imputed pixels.

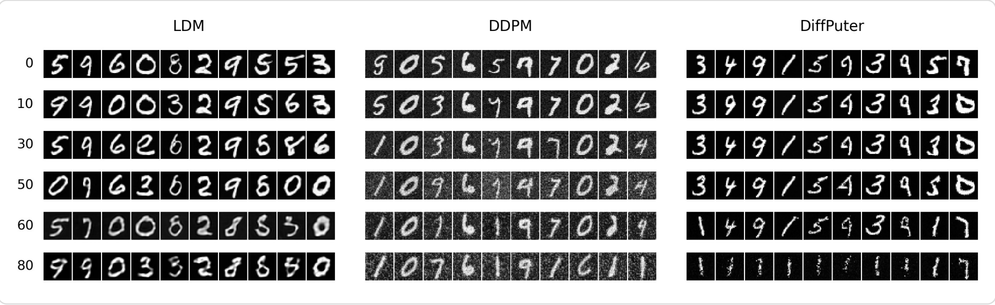

Figure 2. Generated samples across training missing rates. LDMiss stays coherent where DDPM turns grainy.

4 Missing-Data Framework

Each sample splits into observed and missing entries through a binary **MCAR mask**, with missing pixels set to zero. Pixel-space baselines are **MissDiff**, which restricts the loss to observed dimensions, and **DiffPuter**, an EM procedure.

$$\mathcal{L}(\theta) = \mathbb{E}\left[\lambda_t \frac{\|\,\mathbf{m}\odot\left(s_\theta(\mathbf{x}_t,t) - \nabla_{\mathbf{x}_t}\log p_{0t}(\mathbf{x}_t\mid\mathbf{x}_0)\right)\|^2}{\sum_j m_j}\right]$$

For LDMiss the missing dimensions **factor out** under the diagonal-Gaussian encoder.

5 Self-Guided Imputation

Latent space has no notion of missing dimensions, so replacement fails. We condition on the observed pixels. By Bayes' rule the conditional score splits into the unconditional score plus a **guidance term**, our contribution:

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t\mid\mathbf{x}_0^{\text{obs}}) = \nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t) + \nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_0^{\text{obs}}\mid\mathbf{x}_t)$$

The guidance term is approximated with **Tweedie's formula**, decoded to pixel space, and scaled to match the unconditional score. This gives *equal-strength* guidance at every timestep.

6 Experiments

MNIST under MCAR at rates 0, 0.1, 0.3, 0.5, 0.6, and 0.8, with three seeds each. We evaluate sample quality (FID, IS) and imputation (MSE, FID, IS) at a **50% test missing rate.**

Sample Quality

LDMiss retains better FID and IS than pixel-space DDPM at *every* missing rate and stays stable up to 50% missingness. It also shows the lowest variance across seeds. See **Figure 1** below.

Imputation Quality

At a 50% test missing rate, LDMiss stays robust **even when trained with 80% missing data** and beats DDPM guidance and replacement at every rate. DiffPuter matches on MSE at low rates but loses detail at 80%. See **Figure 3** below.

Figure 4. Imputed digits at a 50% test missing rate. Left LDM guidance, center DDPM replacement, right DiffPuter.

Limitations & Future Work

Limitations. Results cover MNIST under MCAR with zero-imputation only.

Future work. Richer datasets, MAR and MNAR mechanisms, and alternative imputation functions.

ACKNOWLEDGEMENTS

Supported by the Villum Foundation (Synergy 50091), the Novo Nordisk Foundation (MLLS, NNF20OC0062606, and DDSA, NNF21SA0069429), the Independent Research Fund Denmark (5334-00122B, 5334-00076B), and the Reinholdt W. Jorck og Hustrus Fond.

KEY RESULTS

LDMiss preserves sample quality to 50% missingness during training and maintains superior imputation performance compared to pixel-space DDM.

> **Why it works.** Diffusing in a learned latent space lets the autoencoder's semantic compression filter zero-imputation artifacts before the score network ever sees them.

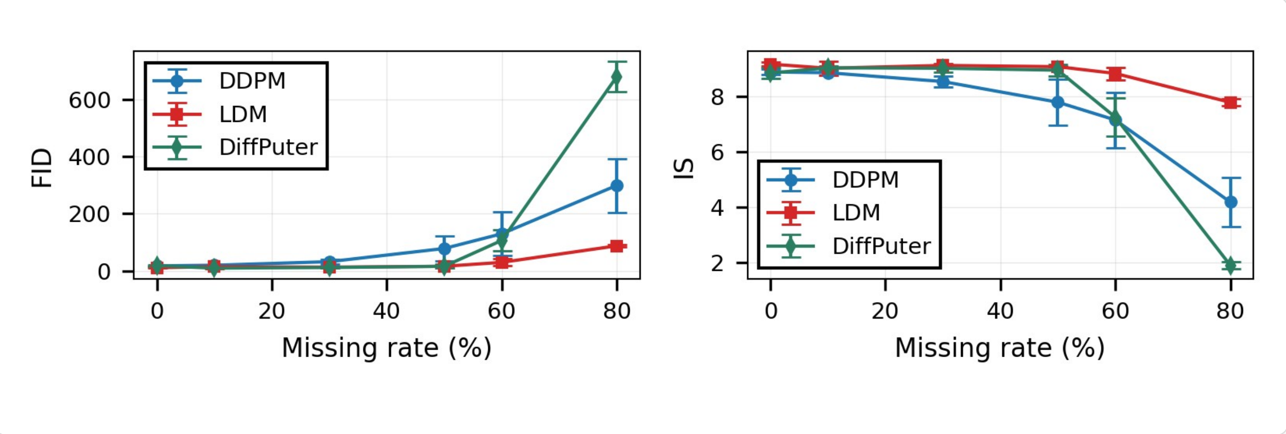

Figure 1. Sample quality (FID down, IS up) versus training missing rate, over three seeds. **LDM (red)** stays flat where DDPM (blue) and DiffPuter (green) rise.

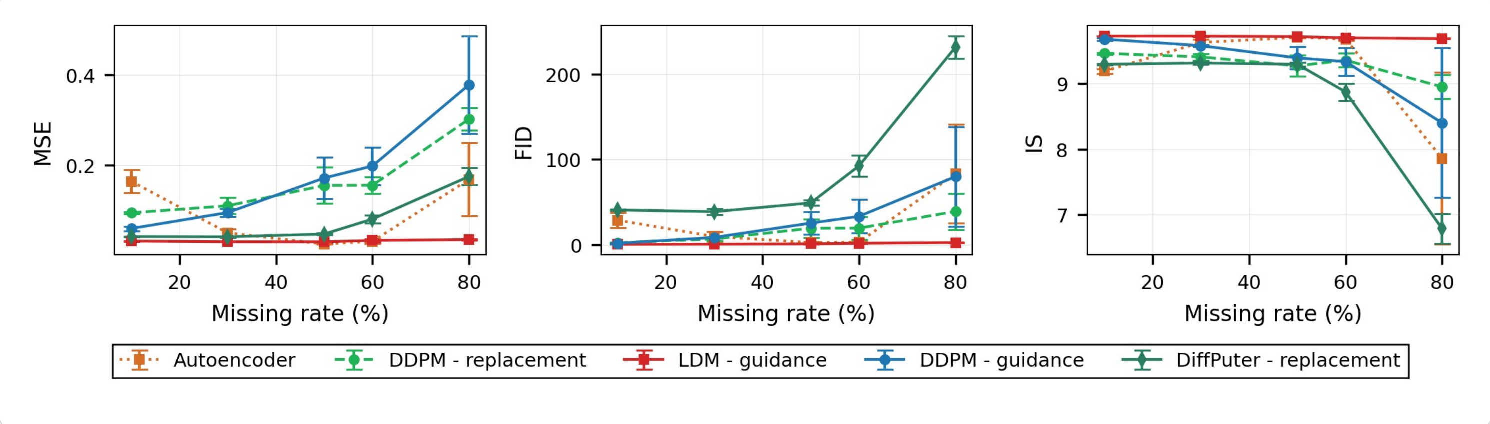

Figure 3. Imputation metrics (MSE, FID, IS) at a 50% test missing rate versus training missing rate. **LDM (red)** stays lowest and flattest.