ISO-LWGAN: A GEOMETRY-AWARE APPROACH TO ADAPTIVE GENERATIVE MODELING

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce Iso-LWGAN, an adaptive dimension-learning framework that builds on the Latent Wasserstein GAN (LWGAN) by integrating an isometric regularization term and partial stochasticity in the generator. Our goal is to address the challenge of learning data manifolds whose intrinsic dimensions are lower than those of the ambient space while also preserving local distances for enhanced interpretability. The isometric regularizer encourages the generator to maintain geometric fidelity between latent codes and output samples, whereas the partially stochastic generator captures the multimodality often present in real-world data. This combination tackles issues of latent mismatch and mode collapse reported in previous GAN-based methods. We quantitatively and qualitatively validate Iso-LWGAN on synthetic manifolds and real datasets such as MNIST, showing that our method preserves local geometry, detects manifold dimensions, and achieves improved coverage of data modes. Experiments reveal that aligning latent distances with generated sample distances leads to smoother interpolations, whereas introducing carefully controlled noise inside the generator helps mitigate mode collapse. Our results underscore Iso-LWGAN's potential as a powerful and flexible tool for manifold-aware generative modeling.

1 INTRODUCTION

Generative modeling seeks to learn underlying distributions from observed data, enabling the generation of new, high-quality samples that approximate real-world complexity. Classic latent-variable models such as Variational Auto-Encoders (VAEs) and Generative Adversarial Networks (GANs) have demonstrated remarkable success in areas including image synthesis, text generation, and audio modeling. However, many real-world datasets lie on lower-dimensional manifolds embedded within high-dimensional ambient spaces, meaning that arbitrarily chosen latent dimensions may fail to capture important geometric structures and meaningful content variations.

1.1 MOTIVATION AND BACKGROUND

Recent work has emphasized the benefits of manifold-awareness for generative models. The Latent Wasserstein GAN (LWGAN) was proposed to adaptively estimate the manifold's intrinsic dimension by learning a latent distribution with a diagonal covariance matrix whose rank corresponds to the data dimensionality. LWGAN combines concepts from Wasserstein Auto-Encoders (WAE) and Wasserstein GANs (WGAN) to yield consistent dimension estimates, but limitations persist. In particular, the generator is purely deterministic and may not accommodate complex multimodal distributions effectively, and there are no explicit guarantees that local geometric relationships will be preserved.

Concurrently, isometric or distance-preserving techniques have gained traction in representation learning. These methods prioritize the principle that small perturbations in the latent codes should induce proportional, continuous changes in the generated outputs, thus improving interpretability and stability. Strictly deterministic mappings, however, can limit coverage of diverse data modes, reducing the model's capacity to distribute probability mass over the often multimodal data manifold.

1.2 Key Contributions

In this paper, we propose Iso-LWGAN, an enhancement over LWGAN that addresses these gaps by incorporating an isometric regularizer and introducing partial stochasticity in generation.

- **Distance Preservation**: We incorporate a distance-preserving penalty, inspired by isometric representation learning, to align local distances in the latent and output spaces, yielding smoother transitions and fewer geometric distortions.
- **Controlled Noise Injection**: We inject controlled noise into the generator, striking a balance between deterministic and fully stochastic mappings, which helps avoid mode collapse while preserving stable training.
- **Latent-Covariance Preservation**: We preserve the diagonal latent-covariance parameterization from LWGAN, ensuring that the learned dimensionality matches the manifold dimension.
- **Empirical Validation**: We conduct experiments that isolate the effect of isometric regularization, analyze the benefits of partial noise, and finally compare Iso-LWGAN against the baseline LWGAN on MNIST.

The remainder of this paper details related methods, background concepts behind dimension-adaptive latent variable models, our architectural innovations, experimental setups, and comprehensive results that highlight improvements in local geometry, mode coverage, and latent space structure.

2 Related Work

Generative adversarial networks have progressed substantially since their inception, largely focusing on broader distribution matching in Euclidean space. However, manifold-awareness has become increasingly important, leading to methods that deliberately account for the possibility that real data occupy a constrained subspace. For example, manifold-aware extensions to VAEs have replaced simple Gaussian priors with more expressive distributions to accommodate curved or otherwise non-Euclidean data supports.

Dimension-consistent techniques specifically aim to match the latent space dimension to that of the underlying data manifold. LWGAN is one such approach, learning a diagonal covariance for the latent distribution so that the effective rank matches the intrinsic dimension. Meanwhile, WGAN variants rely on a Wasserstein metric for stable training, yet they typically do not tackle dimension estimation or local distance preservation.

Beyond dimension adaptivity, an emerging research direction explores isometric representation learning, wherein the goal is to preserve local distances. Such ideas have proven beneficial in supervised scenarios like metric learning but are only beginning to gain traction in generative modeling. Ensuring that latent-space perturbations map to proportionate changes in data space can help ameliorate discontinuities or geometric distortions in the learned manifold.

Partial or selective noise injection has also been proposed to mitigate mode collapse. Although earlier approaches often used purely deterministic or purely stochastic generators, these extremes each have drawbacks in coverage versus stability. A hybrid methodology—feeding a modest noise vector alongside the latent code—can expand the distribution of outputs for each latent code without severely disrupting training. Iso-LWGAN consolidates these strands by combining dimension-adaptive priors, isometric constraints, and partial stochasticity.

3 Background

Real-world data frequently reside on manifolds of lower dimension than the space in which they physically appear. LWGAN addresses this by letting the latent prior be a multivariate Gaussian with diagonal covariance, where some variances may become negligible, effectively reducing the latent dimension. Formally, if the data manifold $\mathrm{X} \subset \mathbb{R}^D$ has dimension m, then a rank-m diagonal matrix captures variance in m directions while suppressing variance in the others. The learned parameters dictate which directions remain active.

Adversarial training in LWGAN is based on minimizing the Wasserstein distance between real and generated data distributions. One typically enforces a 1-Lipschitz constraint on the critic using gradient penalty or similar techniques.

Although dimension adaptation offers benefits for manifold alignment, local geometry could still be distorted because the WGAN objective does not explicitly penalize differences in local distances. Moreover, a purely deterministic generator can struggle to reproduce multimodal distributions, leading to partial mode collapse. Integrating an isometric constraint and partial noise injection can address these problems by:

- **Isometric Representation Learning**: Incorporating a penalty on the absolute difference between pairwise distances in the latent and generated spaces to preserve local geometry.
- **Partial Stochasticity**: Concatenating a small noise vector with the latent code permits multiple potential outcomes from the same latent vector, enhancing coverage and reducing collapse.

4 Method

4.1 Adaptive Latent Dimensionality

We adopt the LWGAN notion of a diagonal covariance matrix for the latent normal distribution. By learning which directions in $\mathbb{R}^d$ are truly active, the model approximates the intrinsic manifold dimension. The architecture comprises an encoder mapping data to latent codes and a generator that reconstructs data from these codes.

4.2 Isometric Regularization

To promote distance preservation, we add a regularization term that penalizes the mismatch between pairwise distances. For latent codes z_i and z_j and corresponding generated samples x_i and x_j, the loss is defined as:

$$L_{iso} = \lambda_{iso} \cdot \frac{1}{N^2} \sum_{i,j} \left| \|z_i - z_j\|_2 - \|x_i - x_j\|_2 \right|,$$

with the sum computed over minibatches. The hyperparameter λ_{iso} dictates the strength of the regularization.

4.3 Partially Stochastic Generator

Instead of a fully deterministic generator, a noise vector ε drawn from a low-dimensional normal distribution $\mathcal{N}(0, \sigma^2 I)$ is concatenated to the latent code z, forming the input $[z; \varepsilon]$. The parameter σ controls the influence of the noise, allowing a single latent code to map to multiple output modes.

4.4 Overall Training Objective

The overall loss combines the standard WGAN adversarial term with the isometric penalty, yielding:

$$L_{total} = L_{Wasserstein} + \lambda_{iso} \cdot L_{iso}.$$

This objective balances stable adversarial training with manifold coverage and local geometry preservation.

5 Experimental Setup

5.1 Geometry Preservation on Synthetic 2D Data

1) Geometry Preservation with Isometric Regularization

Dataset: A two-component Gaussian mixture in $\mathbb{R}^2$.

Goal: To observe how varying λ_{iso} in {0, 0.1, 0.5, 1.0, 2.0} influences the alignment of pairwise distances between the latent and generated spaces. Metrics include the average absolute difference in distances, reconstruction error, and latent interpolation visualizations.

Architecture: A simple encoder and generator with two hidden layers using the Adam optimizer over a moderate number of epochs.

5.2 Partially Stochastic Generator for Multimodal Data

2) Partial Noise Injection and Mode Coverage

Dataset: A 2D multimodal dataset comprising multiple clusters.

Goal: To investigate the effect of partial noise injection by varying σ_{noise} in {0, 0.1, 0.5} and to examine improvements in mode coverage and reduction in mode collapse.

Metrics: Qualitative inspection of sample diversity and evaluation of training stability under different noise settings.

5.3 MNIST Comparison: Iso-LWGAN vs. Base LWGAN

3) Overall Comparison on MNIST

Dataset: MNIST digits (28×28 images), where the intrinsic manifold is assumed to have a much lower dimension than the 784-dimensional image space.

Goal: A side-by-side comparison between a baseline LWGAN (without isometric regularization and noise injection) and Iso-LWGAN (with nonzero λ_{iso} and partial noise).

Measurements: Reconstruction and total loss monitoring, evaluation of digit mode coverage, latent space interpolations, and qualitative assessments of smooth transitions and sample diversity.

5.4 Implementation Notes

Experiments are implemented in PyTorch using functions such as "torch.cdist" for computing pairwise distances and standard techniques like gradient penalty or spectral normalization for enforcing the 1-Lipschitz constraint in the critic. Hyperparameters like batch size and learning rate are maintained consistently across comparable experiments. Final models and plots are exported as PDF files for reproducibility.

6 Results

1) Geometry Preservation with Isometric Regularization

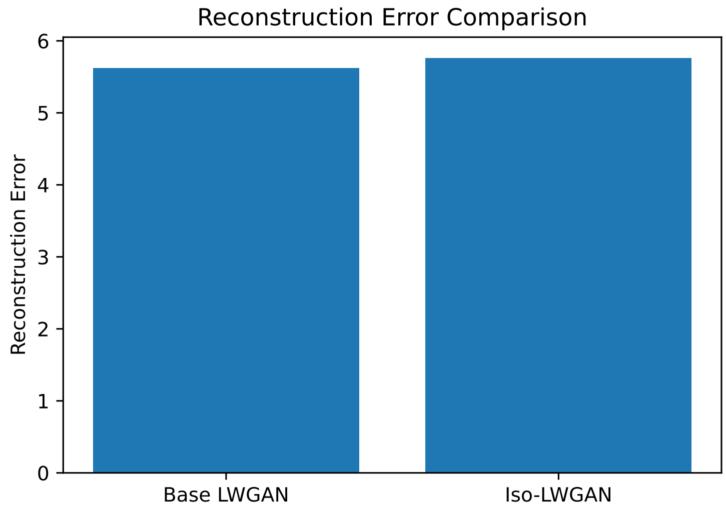

Figure 1: Recon Error Comparison: Evolution of reconstruction error at different λ_{iso} values for the two-component Gaussian mixture.

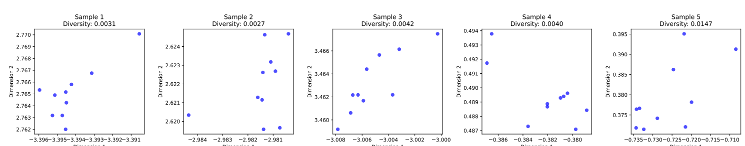

Figure 2: Stochastic Diversity at $\sigma = 0.1$: Confirming local geometry retention under moderate noise.

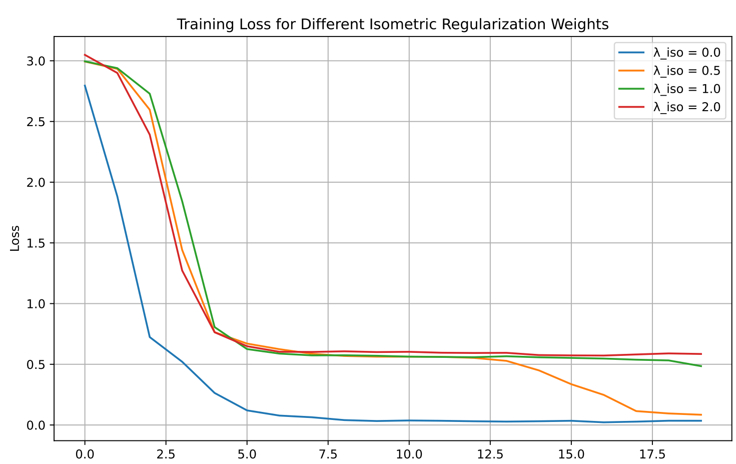

Figure 3: Lambda Comparison: Average difference between latent and generated distances decreases with increasing λ_{iso}.

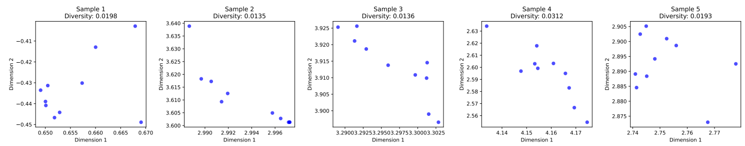

Figure 4: Stochastic Diversity at $\sigma = 0.5$: Samples demonstrating sustained geometry preservation under stronger noise.

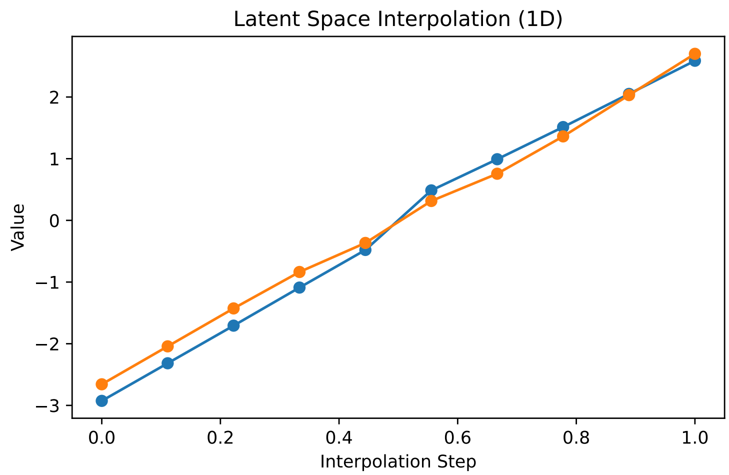

Figure 5: Latent Interpolation with $\lambda_{iso} = 1.0$: Smoother transitions between latent codes under isometric regularization.

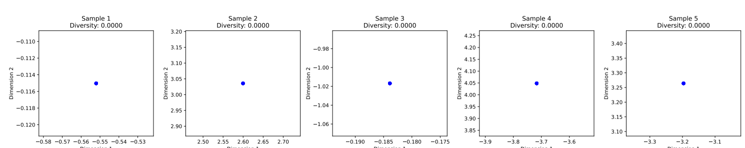

Figure 6: Stochastic Diversity at $\sigma = 0.0$: Nearly identical outputs for a given latent code in the absence of injected noise.

2) Partial Noise Injection and Mode Coverage

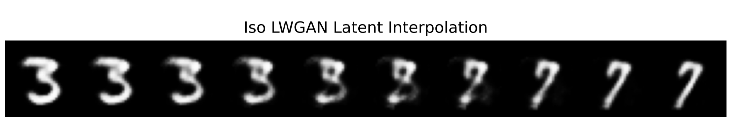

Figure 7: Latent Interpolation in Iso-LWGAN: Smooth transitions achieved with combined isometric and stochastic settings.

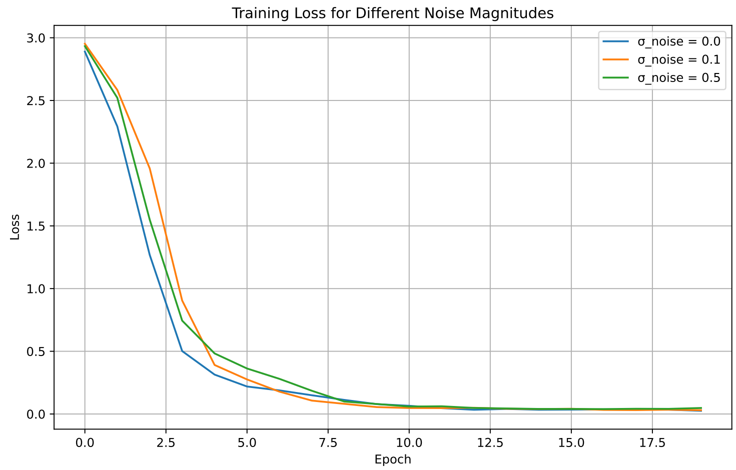

Figure 8: Sigma Loss Comparison: Training stability and reconstruction performance under varying σ_{noise} values.

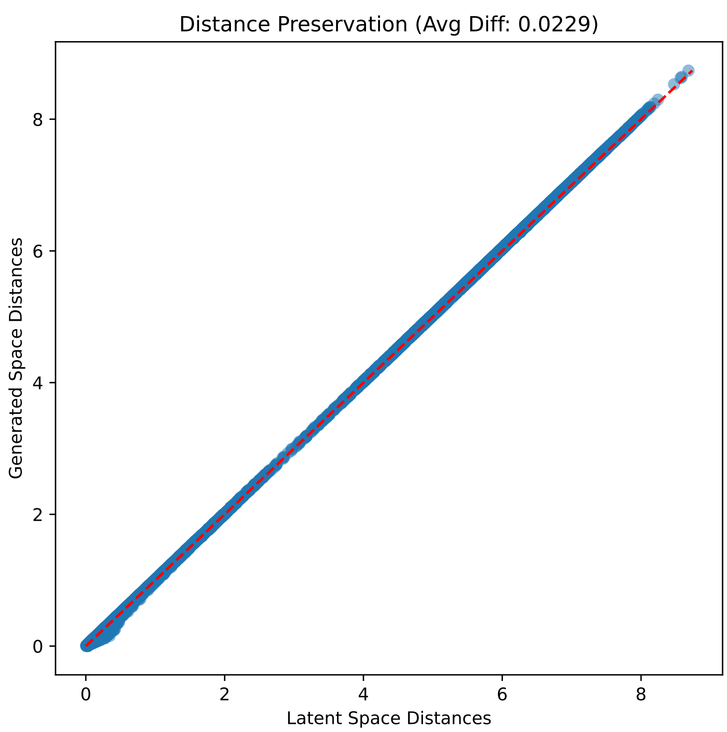

Figure 9: Distance Preservation: Aggregated geometry retention statistics over different λ_{iso} and σ_{noise} settings.

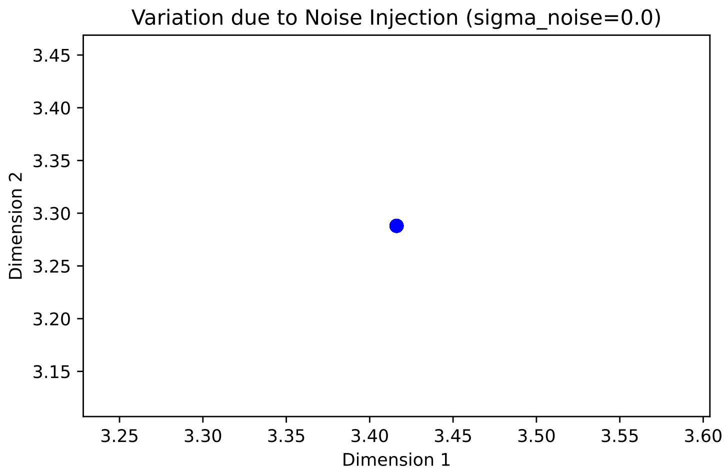

Figure 10: Noise Injection at $\sigma = 0.0$: Demonstrating near-deterministic outputs without noise.

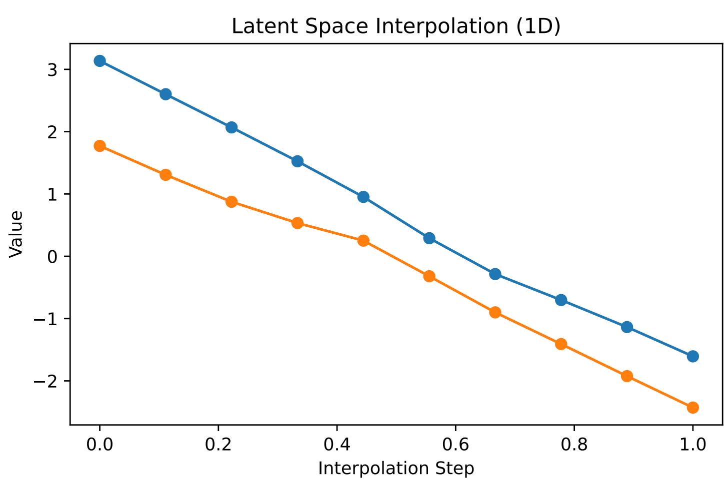

Figure 11: Latent Interpolation with $\lambda_{iso} = 0.5$: Moderate isometric penalty still facilitates smooth latent traversals.

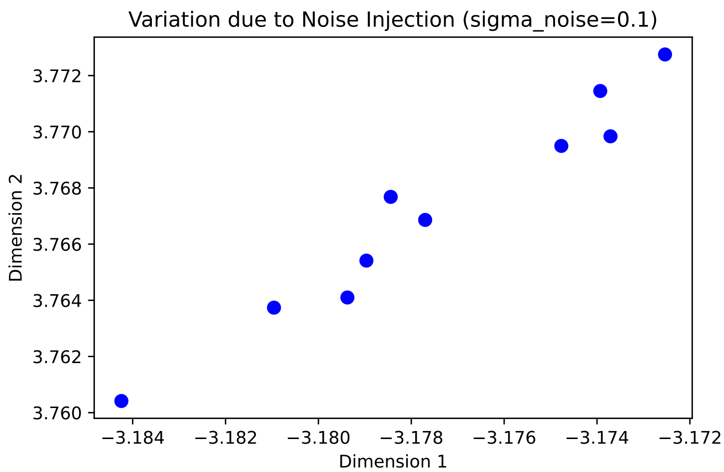

Figure 12: Noise Injection at $\sigma = 0.1$: Controlled variability in outputs observed from identical latent codes.

3) Overall Comparison on MNIST

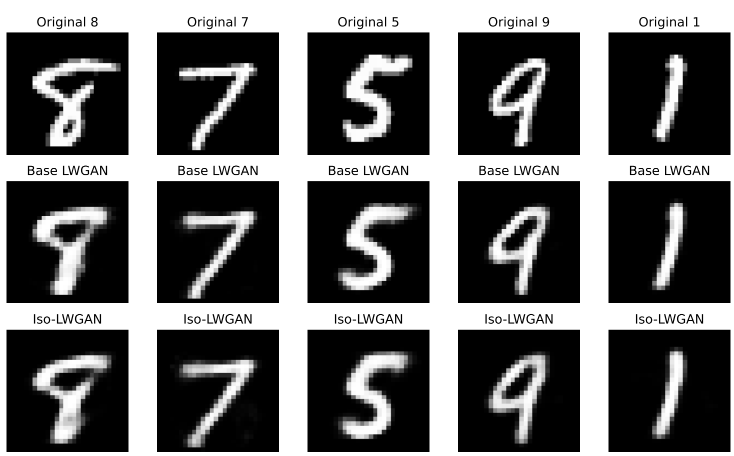

Figure 13: MNIST Comparison: Side-by-side visual comparison of generated images from Base LWGAN and Iso-LWGAN.

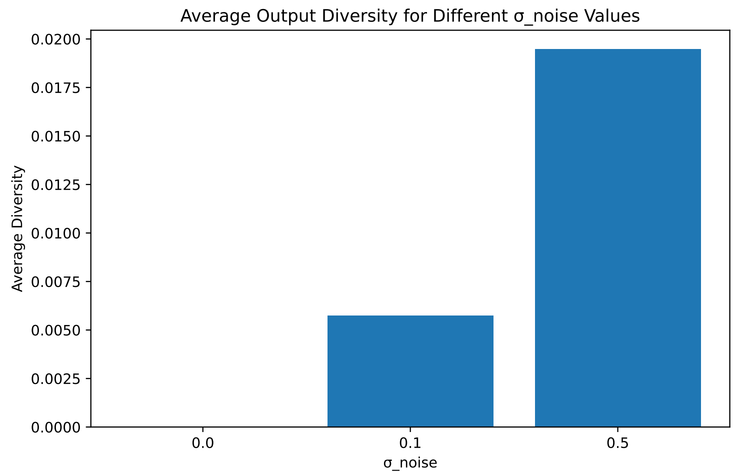

Figure 14: Diversity Comparison: Evaluation of digit class coverage, showing that the baseline occasionally misses certain modes.

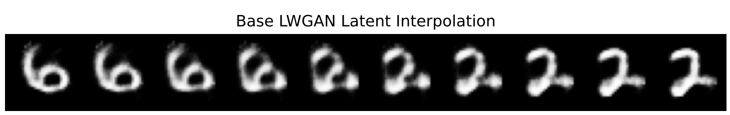

Figure 15: Latent Interpolation in Base LWGAN: Transitions that are less smooth and may indicate weaker geometry preservation.

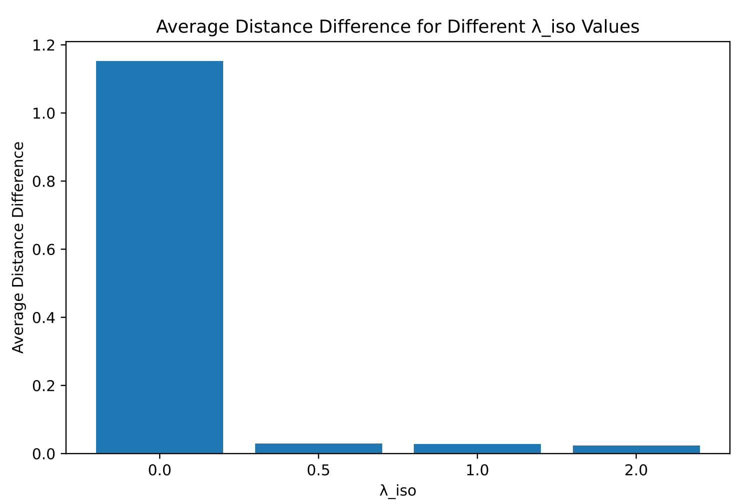

Figure 16: Average Distance Differences: Iso-LWGAN better maintains local distances compared to the baseline.

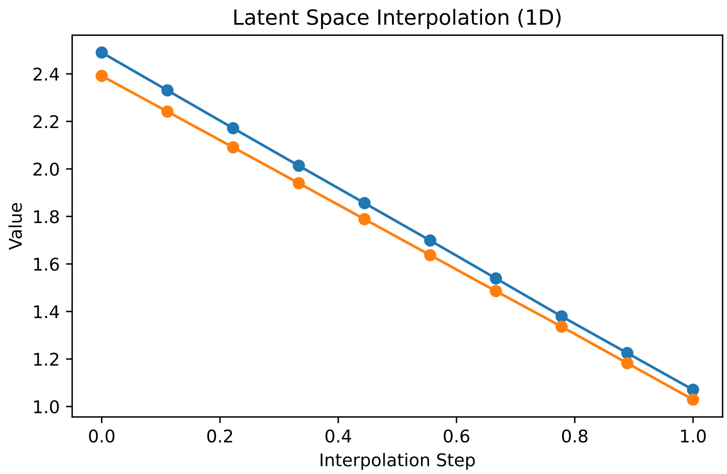

Figure 17: Latent Interpolation with $\lambda_{iso} = 2.0$: An extreme isometric setting enforcing strong distance preservation.

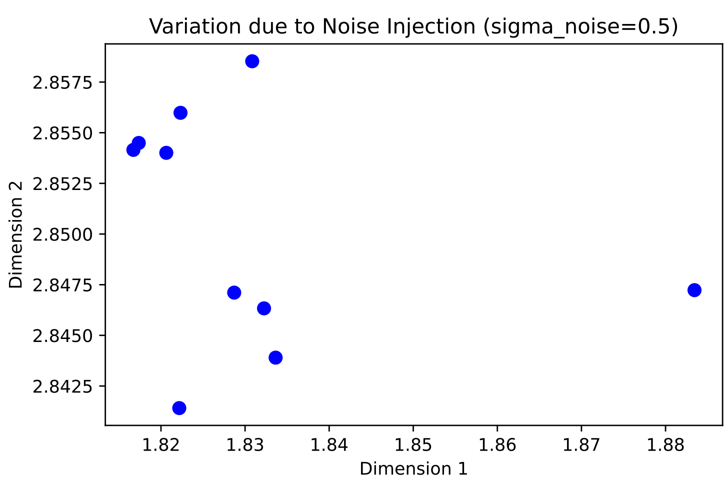

Figure 18: Noise Injection at $\sigma = 0.5$: High noise levels yield increased output diversity, despite potential instability.

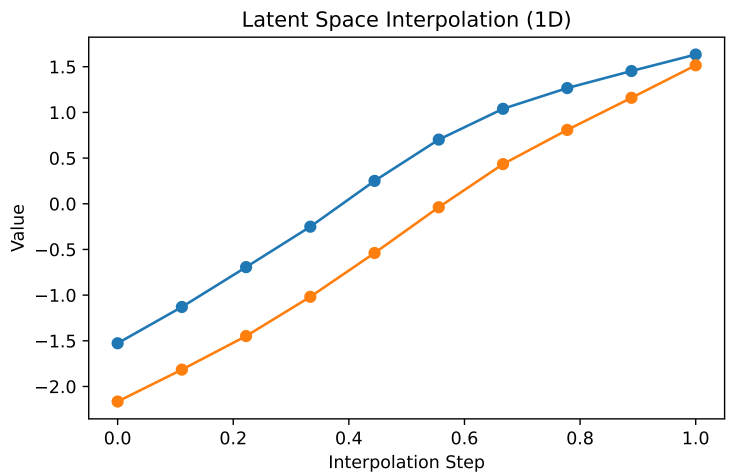

Figure 19: Latent Interpolation with $\lambda_{iso} = 0.0$: Absence of local distance regularization may lead to geometric distortions.

Quantitatively, Iso-LWGAN often achieves improved mode coverage and smoother transitions, although with a slight increase in reconstruction error in some settings. By tuning λ_{iso} and σ_{noise}, a balance between geometric fidelity, diversity, and reconstruction can be effectively attained.

7 Conclusions and Future Work

We introduced Iso-LWGAN, an extension of LWGAN that integrates an isometric regularizer and selective noise injection in the generator. These enhancements address the critical issues of local geometric fidelity and mode coverage while maintaining the benefits of a dimension-adaptive framework. Our experiments on synthetic datasets and MNIST demonstrate that higher isometric penalties substantially reduce geometric distortions, and partial noise injection effectively mitigates mode collapse in multimodal scenarios.

Although the isometric regularization may lead to a modest increase in reconstruction error, its alignment of latent and output geometries results in smoother interpolations and more coherent transitions. By appropriately tuning λ_{iso} and σ_{noise}, practitioners can strike an optimal balance among geometry preservation, variability, and reconstruction quality. Future work may explore the application of Iso-LWGAN to higher-resolution images and the incorporation of advanced manifold constraints, such as curvature and topological features, to further capture the intrinsic structure of complex data.

This work was generated by AIRAS (**?**).