EXTENDABLE AND INVERTIBLE MANIFOLD LEARNING WITH GEOMETRY REGULARIZED AUTOENCODERS

Andrés F. Duque[*], *Sacha Morin*[**†], *Guy Wolf* [**†] *and Kevin R. Moon*[*]

[*]Utah State University, [**]Université de Montréal, [†]Mila – Quebec AI Institute
Research partially funded by IVADO – Institut de Valorisation des Données

Introduction

A fundamental task in data exploration is to extract simplified low dimensional representations for faithfully visualizing data in two or three dimensions. Common approaches to this task include kernel methods for manifold learning and autoencoders, which have the following weaknesses:

- **Kernel methods** typically lack out-of-sample extensions and an inverse mapping from the latent representation to the ambient space.
- **Autoencoders** (AE) usually do not capture the global intrinsic geometry present in the data.

We integrate both approaches by incorporating a geometric regularization term in the bottleneck of the autoencoder, resulting in the following contributions:

Our contributions:

- Improve AE latent representation for visualizing the true semantics of the data.
- Improve reconstruction error and disentanglement of the latent variables.
- Provide manifold learning methods with an invertible mapping and out-of-sample extension.

Geometric Regularization

Denote an embedding $\mathcal{E}_p = \{e_1, e_2, \ldots, e_n\}$ produced by a manifold learning method, where $e_i \in \mathbb{R}^d$ is the embedding coordinates of each observation i. The geometry regularized autoencoder (**GRAE**) is trained using eq. (1).
We use PHATE [4] to learn $\mathcal{E}$, but GRAE can be easily implemented using other methods e.g. UMAP.

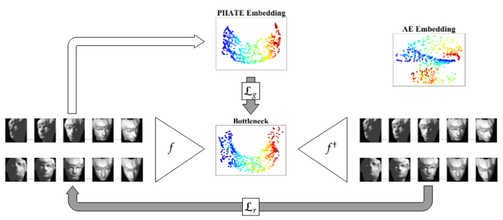

Figure 1: **Overview of GRAE** on the Faces dataset [6]. Geometric regularization is applied to enforce similarity between GRAE and PHATE embeddings. The vanilla AE embedding (top right) is added for reference.

$$\underset{f,f^\dagger}{\arg\min}\, \mathcal{L}(f, f^\dagger) = \mathcal{L}_r(X, f^\dagger(f(X)) + \lambda \mathcal{L}_g(f(X), \mathcal{E}). \quad (1)$$

Where $f : \mathcal{M} \to \mathbb{R}^d$ and $f^\dagger : \mathbb{R}^d \to \mathcal{M}$

- Autoencoders rely solely on the reconstruction loss $\mathcal{L}_r$
- We introduce a geometric loss $\mathcal{L}_g(f(X), \mathcal{E}) = \sum_{i=1}^{n} \|e_i - f(x_i)\|^2$

Qualitative Results

We experimentally compare GRAE with a standard AE, Embedding with Autoencoder Regularization (EAER) [8], Diffusion Nets [3], Topological Autoencoders (TAE) [5], and UMAP [2] on six benchmark datasets: **1)** The "Swiss Roll" with a thin middle slice removed from the training set and used for testing. **3)** Three MNIST digits picked randomly and rotated 360 times at one-degree intervals, for a total of 1080 images. **3)** Teapot [7]. **4)** UMIST Faces dataset [1]. **5)** Object tracking synthetic dataset. **6)** Single-cell mass cytometry data showing iPSC reprogramming of mouse embryonic fibroblasts (hereinafter, "iPSC") as introduced in [9].

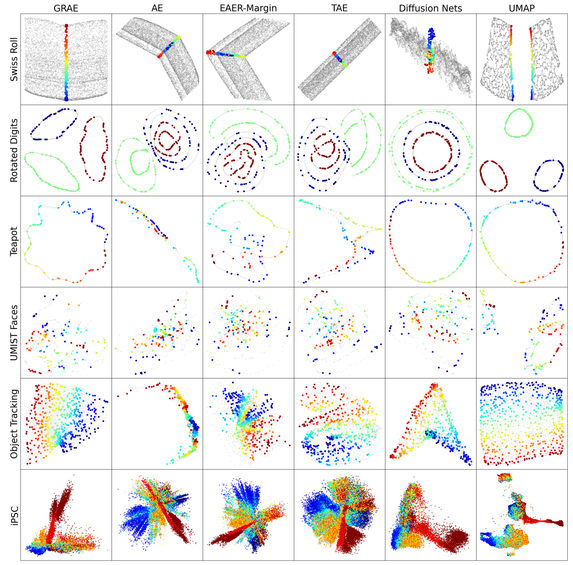

Figure 2: **Latent representations** from all considered methods on the benchmark datasets

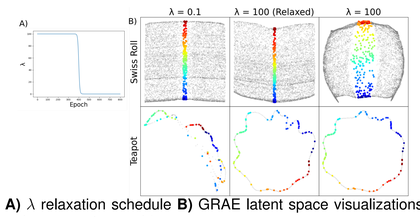

Figure 3: **A)** λ relaxation schedule **B)** GRAE latent space visualizations with three different λ schemes on two problems

- We first notice that GRAE recovers a sensible geometry for all problems while other methods fail at basic tasks such as uncoiling the Swiss Roll, disentangling the Rotated Digits, showing the coordinate plane on Object Tracking or recovering a circular structure on Teapot. Only GRAE and Diffusion Nets show the two expected branches of the iPSC manifold. UMAP also shows good geometry in general, but tears the overall structure of the manifold on the Swiss Roll and iPSC problems despite their known continuous nature.
- The λ schedule during training in Figure 3 helps to find a better reconstruction after a good latent representation has been found.

Quantitative Results

Dataset	Model	Metrics		
		R^2	*MSE*	*Rel. MSE*
Swiss Roll	GRAE	**0.9762 (1)**	**0.0034 (1)**	**-83.8 % (1)**
	AE	0.5136 (3)	0.0210 (4)	0.0 % (4)
	EAER-Margin	0.3746 (6)	0.0246 (5)	17.1 % (5)
	TAE	0.4905 (4)	0.0196 (3)	-6.7 % (3)
	Diffusion Nets	0.4257 (5)	0.0546 (6)	160.0 % (6)
	UMAP	**0.8455 (2)**	**0.0042 (2)**	**-80.0 % (2)**
Rotated Digits	GRAE	**0.9829 (2)**	**0.0027 (1)**	**-69.3 % (1)**
	AE	0.3757 (3)	0.0088 (3)	0.0 % (3)
	EAER-Margin	0.3181 (5)	**0.0061 (2)**	**-30.7 % (2)**
	TAE	0.3359 (4)	0.0101 (4)	14.8 % (4)
	Diffusion Nets	0.2530 (6)	0.0300 (5)	240.9 % (5)
	UMAP	**0.9845 (1)**	0.0653 (6)	642.0 % (6)
Teapot	GRAE	**0.9989 (1)**	**0.0032 (2)**	**-62.4 % (2)**
	AE	0.2079 (6)	0.0085 (4)	0.0 % (4)
	EAER-Margin	0.2526 (4)	**0.0027 (1)**	**-68.2 % (1)**
	TAE	0.2287 (5)	0.0097 (5)	14.1 % (5)
	Diffusion Nets	0.9933 (3)	0.0038 (3)	-55.3 % (3)
	UMAP	**0.9981 (2)**	0.0160 (6)	88.2 % (6)
UMIST Faces	GRAE	0.9371 (3)	**0.0092 (1)**	**-35.7 % (1)**
	AE	0.9040 (6)	0.0143 (5)	0.0 % (5)
	EAER-Margin	0.9298 (5)	**0.0108 (2)**	**-24.5 % (2)**
	TAE	**0.9426 (1)**	0.0118 (3)	-17.5 % (3)
	Diffusion Nets	**0.9407 (2)**	0.0128 (4)	-10.5 % (4)
	UMAP	0.9348 (4)	0.0292 (6)	104.2 % (6)
Object Tracking	GRAE	**0.9298 (2)**	**0.0410 (1)**	**-6.6 % (1)**
	AE	0.3658 (6)	0.0439 (5)	0.0 % (5)
	EAER-Margin	0.4124 (5)	**0.0429 (2)**	**-2.3 % (2)**
	TAE	0.4369 (4)	0.0434 (3)	-1.1 % (3)
	Diffusion Nets	0.7806 (3)	0.0435 (4)	-0.9 % (4)
	UMAP	**0.9890 (1)**	0.0855 (6)	94.8 % (6)
iPSC	GRAE	**0.2620 (2)**	**0.7440 (2)**	**1.0 % (2)**
	AE	0.0919 (6)	**0.7366 (1)**	**0.0 % (1)**
	EAER-Margin	0.1305 (4)	0.7721 (4)	4.8 % (4)
	TAE	0.1296 (5)	0.7603 (3)	3.2 % (3)
	Diffusion Nets	0.2571 (3)	1.1060 (6)	50.1 % (6)
	UMAP	**0.3609 (1)**	0.7741 (5)	5.1 % (5)

Above we report a quantitive assessment of all considered methods. We score models on two tasks: i) recovering ground truth factors in a meaningful way in the latent space and ii) reconstructing samples from the latent space back to the input space adequately.
Reconstruction quality is assessed using the MSE between the input and the reconstruction. Disentanglement of the latent factors is assessed by fitting a linear regression to predict said factors, using the embedding coordinates as regressors.

- GRAE comes first on reconstruction quality in most of the datasets. The addition of the soft contraint in eq. 1 actually can help to find a better reconstruction, guiding the AE during training to a better latent representation.
- The high correlation between the GRAE embedding coordinates and the ground truth latent coordinates show disentanglement potential.

Scalability of PHATE

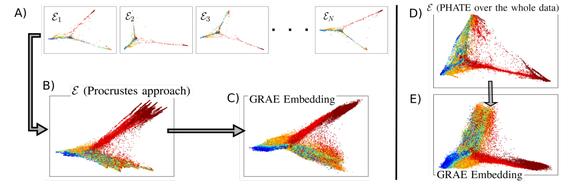

Figure 4: **Scalable GRAE.** Overview of GRAE applied to 200,000 observations of iPSC data. **A)** Multiple mini-batch embeddings, each of which share some common observations. **B)** Combined embedding using the Procrustes method to align the mini-batch embeddings in a consistent way. **C)** GRAE embedding using (**B**) as $\mathcal{E}$ in the geometric loss. **D)** PHATE embedding computed over the whole data set. **E)** GRAE embedding using (**D**) as $\mathcal{E}$ in the geometric regularization. Although both approaches produce near similar embeddings, the mini-batch approach is scalable. Thus, GRAE can tackle bigger datasets than PHATE

Impact of Geometric Regularization on Reconstruction Quality

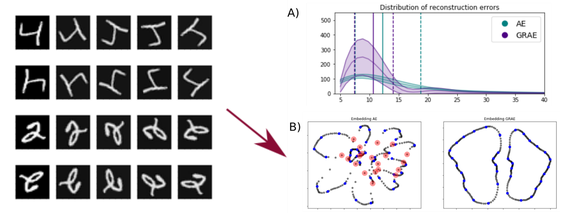

Figure 5: **A)** Distributions of errors of two rotated digits averaged over ten runs for AE vs GRAE. Dashed lines represent the 1st and 3rd quartiles, and solid lines represent the median. We notice that AE is more unstable than GRAE, having a heavier tail, since it fails completely to reconstruct certain images, while GRAE typically presents lighter tails. **B)** Typical embeddings produced for AE and GRAE. Blue points represent a subsample of the training data (subsampling only done for visualization purposes). Black points are the generated points on the latent space via interpolation. Red colored points in the AE embedding represent the 20 interpolated points with the highest reconstruction error. We observe that bad reconstruction typically occurs in sparse regions or crossing lines, i.e., in regions with poorly learned geometry.

References

[1] Daniel B Graham and Nigel M Allinson. "Characterising virtual eigensignatures for general purpose face recognition". In: *Face Recognition*. Springer, 1998, pp. 446–456.

[2] Leland McInnes, John Healy, and James Melville. "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction". In: *arXiv e-prints*, arXiv:1802.03426 (Feb. 2018). arXiv: 1802.03426 [stat.ML].

[3] Gal Mishne et al. "Diffusion nets". In: *Applied and Computational Harmonic Analysis* 47.2 (2019), pp. 259–285.

[4] Kevin R. Moon et al. "Visualizing structure and transitions in high-dimensional biological data". In: *Nature Biotechnology* 37.12 (2019), pp. 1482–1492.

[5] Michael Moor et al. "Topological Autoencoders". In: *arXiv preprint arXiv:1906.00722* (2019).

[6] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. "A global geometric framework for nonlinear dimensionality reduction". In: *science* 290.5500 (2000), pp. 2319–2323.

[7] Kilian Q Weinberger, Fei Sha, and Lawrence K Saul. "Learning a kernel matrix for nonlinear dimensionality reduction". In: *Proceedings ICML*. 2004, p. 106.

[8] Wenchao Yu et al. "Embedding with autoencoder regularization". In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer. 2013, pp. 208–223.

[9] Eli R Zunder et al. "A continuous molecular roadmap to iPSC reprogramming through progression analysis of single-cell mass cytometry". In: *Cell stem cell* 16.3 (2015), pp. 323–337.