Witness Autoencoder: Shaping the Latent Space with Witness Complexes

Simon Scönenberger[1,2], Anastasiia Varava[1], Vladislav Polianskii[1], Jen Jen Chung[2], Roland Siegwart[2], Danica Kragic[1]

[1]Robotics, Perception and Learning Group (KTH), [2]Autonomous Systems Lab (ETH Zürich)

NEURAL INFORMATION PROCESSING SYSTEMS

Abstract

We present a *Witness Autoencoder* (W-AE) – an autoencoder that captures geodesic distances of the data in the latent space. Our algorithm uses witness complexes to compute geodesic distance approximations on a mini-batch level and leverages topological information from the entire dataset while performing batch-wise approximations. This way, our method allows to capture the global structure of the data even with a small batch size, which is beneficial for large-scale real-world data. We show that our method captures the structure of the manifold more accurately than the recently introduced topological autoencoder (TopoAE).

Motivation

Currently, autoencoders (AEs) are widely used for non-linear dimensionality reduction in various applications, mainly due to the expressiveness of neural networks and the encoder-decoder architecture. However, one of the key issues of AEs is that their latent spaces do not necessarily reflect the geometric and topological structure of the true data manifold – i.e., they are not guaranteed to preserve relative distances between points and the topological structure of the data. Preserving this structure is beneficial not only for interpretability of the latent space, but also for generalization capabilities [1,2] and robustness to adversarial attacks [3] .

Proposed Method

We present a *Witness Autoencoder* (W-AE), which introduces a novel loss term for autoencoders to enforce structure preservation in the latent space, i.e. alignment of the geodesic distances between data and latent space. Thus the total loss becomes,

$$\mathcal{L}(x) \coloneqq \mathcal{L}_r(x, \hat{x}) + \lambda_t \mathcal{L}_t \,.$$

The loss term is closely related to the ones presented in [10, 11] and relies on the construction of a neighborhood graph to approximate distances on the manifold. For the construction of such a neighborhood graph, we present a new method that is based on witness complexes. It improves geodesic distance approximations on a mini-batch level, which is desirable since false neighbors and missed neighbors are more likely to occur for small mini-batches since low density regions are more likely. In figure 1 our neighborhood graph construction is compared to the method used in TopoAE and k-NN neighborhood graphs for the Swiss roll dataset, where wrong neighbors occur as so called short-circuit errors.

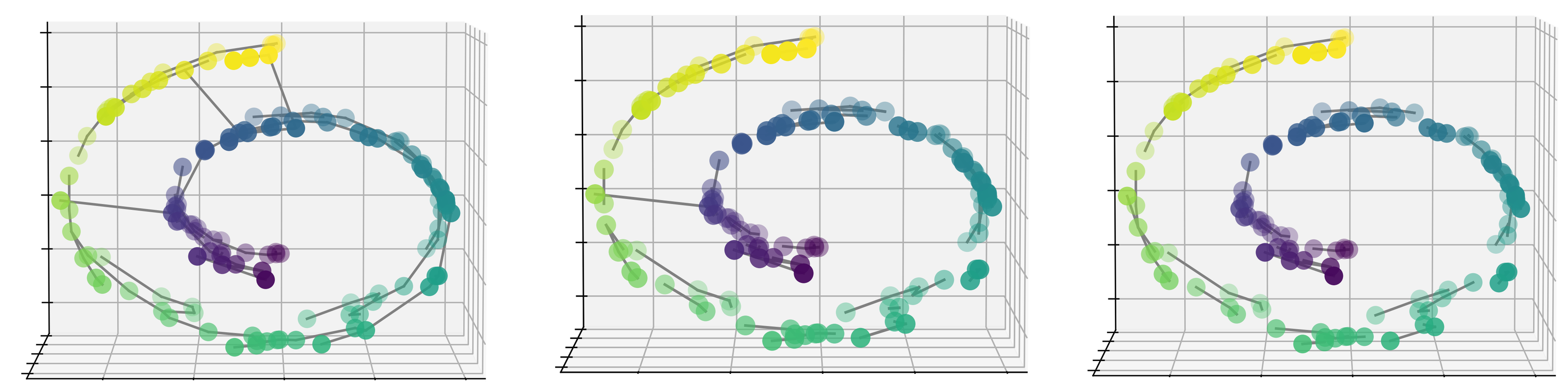

Figure 1: Different graphs constructed on the Swiss roll dataset. The graphs constructed from edges of VR 0-order persistence pairings (left) and 1-NN (middle) fail to approximate the geodesic distances, i.e. there are short-circuit errors. 1-NN-WC (right) approximates the geodesic distances well. ($n_{bs} = 128, |W| = 2048$)

Witnessing a neighborhood graph

To construct a neighborhood graph on a mini-batch level, that leverages topological information of the entire dataset we rely on a witness complex filtration. We refer to the datapoints of the mini-batch as landmark points ($\mathbb{L}$) from which we construct a graph and the entire dataset as witnesses (W), that determine at what *filtration value* R an edge occurs. Formally for a defined distance measure $d(\cdot)$ an edge (l_1, l_2) occurs at filtration value R iff,

$$argmin_{w \in W} \max\big(d(l_1, w), d(l_2, w)\big) = R.$$

The filtration values for all possible pairs in the mini-batch then determine the k-NN. Intuitively, this formulation "shortens" the distance measured along the manifold, since it is likely that a witness lies between two neighbors, whereas a witness between non-neighbors is unlikely. Hence the number of short-circuit errors can be reduced (see figure 1 and table 1).

Table 1: Observed number of mini-batches out of 100 containing short-circuit errors/wrong neighbors for 0-order persistence pairings of a VR filtration (VR), k-NN and k-NN-WC.

n_{bs}	Method	Neighbors (k)			
		N/A	1	2	3
64	VR	100	-	-	-
	k-NN	-	77	100	100
	k-NN-WC	-	5	11	72
128	VR	61	-	-	-
	k-NN	-	35	76	100
	k-NN-WC	-	0	2	10
256	VR	6	-	-	-
	k-NN	-	3	11	77
	k-NN-WC	-	0	0	0

Experimental Study on Swiss Roll Dataset

For evaluation W-AE was train with different batch-sizes on the Swiss roll dataset and compared to TopoAE, UMAP, tSNE and a vanilla AEs. In figure 2 qualitative results are presented and compared to TopoAE, that show that W-AE approximates geodesic distances better even for small mini-batch sizes (n_{bs}). In the publication, qualitative and quantitative results for all methods are presented.

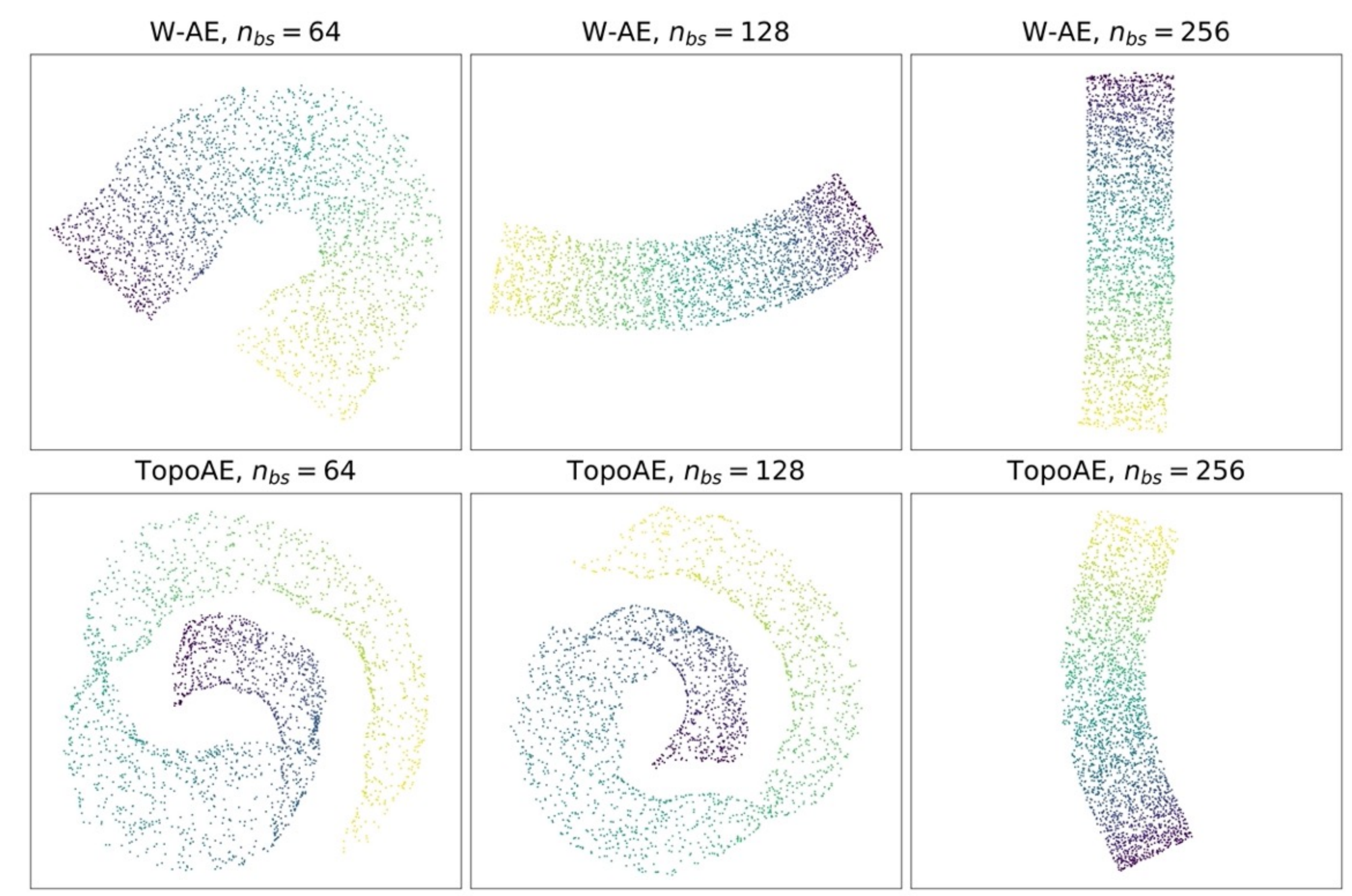

Figure 2: Latent representation obtained with TopoAE and W-AE of Swiss roll dataset for different mini-batch sizes. W-AE is able preserve the structure of the manifold better for smaller batch sizes (n_{bs}).

References

[1] Cem Anil, James Lucas, and Roger Grosse. Sorting out Lipschitz function approximation. *ICML 2019*,

[2] Peter L. Bartlett, Dylan J. Foster, and M. Telgarsky. Spectrally-normalized margin bounds for neural networks Peter, (Nips 2017).

[3] Yusuke Tsuzuku, Issei Sato, and Masashi Sugiyama. Lipschitz-margin training: Scalable certification of perturbation invariance for deep neural networks. NeurIPS 2018

[4] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. Topological Autoencoders. ICML 2019

[5] Stan Z. Li, Zelin Zhang, and Lirong Wu. Markov-Lipschitz Deep Learning. 2020