Information Sheet

What is the main claim of the paper? Why is this an important contribution to the machine learning literature?

The manuscript describes a solution to the problem of visualizing temporal, high-dimensional data in a point-based, two-dimensional visualization. While many nonlinear dimensionality reduction approaches have been proposed [5, 1], only a few of these approaches consider the temporal nature of the data during the embedding construction process, but no one has directly and simultaneously addressed the problem of constraining the optimization to address both temporal constraints and ease of visual explanation. Our approach is important because it addresses both concerns in a straightforward, unified way using additional factors to the t-SNE optimization function.

What papers by other authors make the most closely related contributions, and how is your paper related to them?

In a typical visualization task, high-dimensional data is embedded in two-dimensional space using one of many dimensionality reduction techniques. The points are then visualized using a scatter plot. Arrows are often superimposed on the visualization to show the temporal dependence between data points. However, since most dimensionality reduction techniques do not take into account the temporal aspect of the data, the arrows often appear scattered, and temporal patterns cannot be easily deciphered from the visualization. Existing approaches that do account for temporal dynamics typically construct a sequence of two-dimensional embeddings and stack them along a third time dimension [3, 1]. While attractive, Sedlmair et al. [4] showed that three-dimensional embeddings offer little practical advantage over two-dimensional embeddings, while at the same time often requiring interactive tools for their effective use.

In this work, we propose two loss functions, the *directional coherence loss* (DCL) and *edge length loss* (ELL), to promote the formation of temporally coherent structures in the embedding space. These loss functions can be easily integrated into existing algorithms such as t-SNE and UMAP. These two losses each introduce two additional parameters, and we include parameter studies. We perform an extensive parameter study and quantiative analysis, comparing our approach to existing methods on three synthetic and two real-world data sets.

Have you published parts of your paper before, for instance in a conference? If so, give details of your previous paper(s) and a precise statement detailing how your paper provides a significant contribution beyond the previous paper(s).
This manuscript is an extended version of our paper originally presented at the *Discovery Science 2023* conference, entitled "*Refining Temporal Visualizations Using the Directional Coherence Loss*" [2]. Our submission here represents a substantial extension of the previously published paper: we newly propose the ELL loss, perform extensive parameter studies along with a quantiative analysis of the resulting embeddings, and compare our approach to existing methods in the field.

References

[1] L. McInnes, J. Healy, and J. Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv e-prints*, 2018.

[2] P. G. Poličar and B. Zupan. Refining temporal visualizations using the directional coherence loss. In A. Bifet, A. C. Lorena, R. P. Ribeiro, J. Gama, and P. H. Abreu, editors, *Discovery Science*, pages 204–215, Cham, 2023. Springer Nature Switzerland.

[3] P. E. Rauber, A. X. Falcão, and A. C. Telea. Visualizing time-dependent data using dynamic t-SNE. In *EuroVis 2016 - Short Papers*, page 73–77, Eindhoven, The Netherlands, 2016. The Eurographics Association.

[4] M. Sedlmair, T. Munzner, and M. Tory. Empirical guidance on scatterplot and dimension reduction technique choices. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2634–2643, 2013.

[5] L. Van der Maaten and G. Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(Nov):2579–2605, 2008.