

A manifold fitting approach for high-dimensional data reduction beyond Euclidean space

January 29 2024

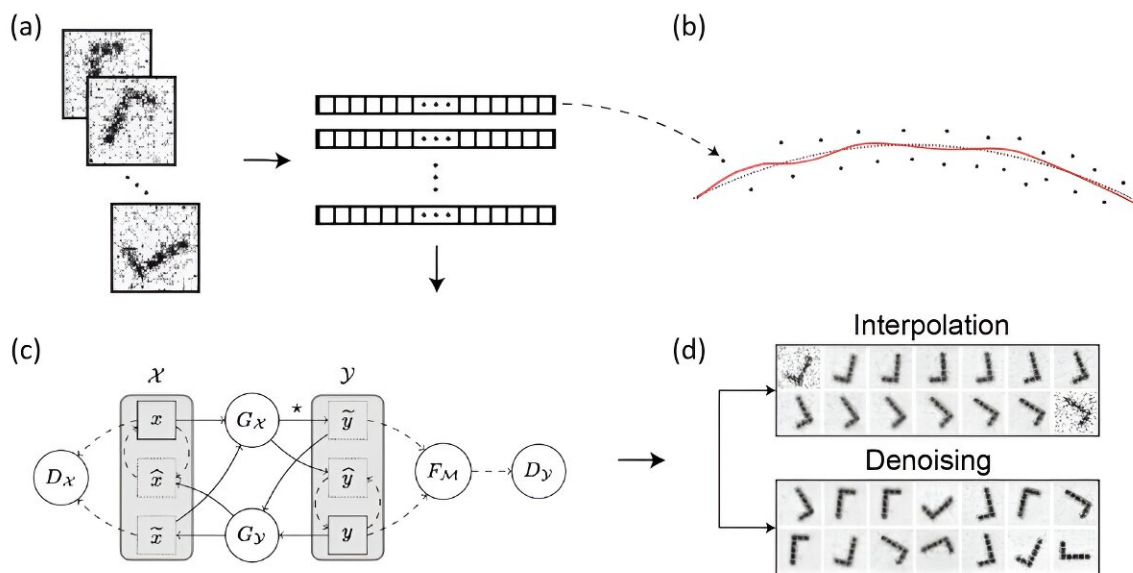


Illustration of fitting the latent manifold using the Cycle Generative Adversarial Network (CycleGAN). CycleGAN is a deep learning technique for unsupervised image-to-image translation. In the real world, data, such as the images shown in panel (a), are often high-dimensional vectors. These vectors typically reside around a low-dimensional latent manifold, depicted by the black dotted curve in panel (b). The CycleGAN framework, detailed in panel (c), effectively learns to estimate this latent manifold (illustrated as the red curve in panel (b)). This advancement facilitates nonlinear interpolation and denoising within the high-dimensional ambient space (panel (d)), offering significant improvements in data

processing and analysis. Credit: *Proceedings of the National Academy of Sciences* (2024). DOI: 10.1073/pnas.2311436121

Statisticians from the National University of Singapore (NUS) have introduced a new technique that accurately describes high-dimensional data using lower-dimensional smooth structures. This innovation marks a significant step forward in addressing the challenges of complex nonlinear dimension reduction.

Traditional data analysis methods often rely on Euclidean (linear) dependencies among features. While this approach simplifies data representation, it struggles to capture the underlying complex patterns in [high-dimensional data](#), typically located close to low-dimensional manifolds.

To bridge this gap, manifold-learning techniques have emerged as a promising solution. However, existing methods, such as manifold embedding and denoising, have been limited by a lack of detailed geometric understanding and robust theoretical underpinnings.

The team, led by Associate Professor Zhigang Yao from the Department of Statistics and Data Science, NUS with his Ph.D. student Jiaji Su pioneered a novel method for effectively estimating low-dimensional manifolds hidden within high-dimensional data. This approach not only achieves cutting-edge estimation accuracy and convergence rates but also enhances computational efficiency through the utilization of deep Generative Adversarial Networks (GANs).

This work was conducted in collaboration with Professor Shing-Tung Yau from the Yau Mathematical Sciences Center (YMSC) at Tsinghua University. Part of the work comes from Prof. Yao's collaboration with

Prof. Yau during his sabbatical visit to the Center of Mathematical Sciences and Applications (CMSA) at Harvard University.

Their findings have been [published](#) as a methodology paper in the *Proceedings of the National Academy of Sciences*.

Prof. Yao delivered a 45-minute invited lecture on this research at the recent [International Congress of Chinese Mathematicians](#) (ICCM) held in Shanghai, Jan. 2–5, 2024.

Highlighting the significance of the work, Prof. Yao said, "By accurately fitting manifolds, we can reduce data dimensionality while preserving crucial information, including the underlying geometric structure. This represents a major leap in data analysis, enhancing both accuracy and efficiency. By providing a solution that overcomes the limitations of previous methods, our research paves the way for enhanced data analysis and offers valuable insights for diverse applications in the scientific community."

Looking ahead, Yao's research team is developing a new framework to process even more complex data, such as single-cell RNA sequence data, while continuing to collaborate with the YMSC team. This ongoing work promises to revolutionize the approach for the reduction and processing of complex datasets, potentially offering new insights into a range of scientific fields.

More information: Zhigang Yao et al, Manifold fitting with CycleGAN, *Proceedings of the National Academy of Sciences* (2024).
[DOI: 10.1073/pnas.2311436121](https://doi.org/10.1073/pnas.2311436121)

Provided by National University of Singapore

Citation: A manifold fitting approach for high-dimensional data reduction beyond Euclidean space (2024, January 29) retrieved 2 May 2024 from <https://phys.org/news/2024-01-manifold-approach-high-dimensional-reduction.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.