

A Survey of Manifold Learning and Its Applications for Multimedia

Hannes Fassold

Joanneum Research – Digital, Graz, Austria

Email: Hannes.fassold@joanneum.at (H.F.)

Abstract—Manifold learning is an emerging research domain of machine learning and helps to overcome the limits of standard neural networks which are restricted to Euclidean spaces. In this work, we give an introduction into manifold learning and how it is employed for important application fields in multimedia like similarity search, image classification, synthesis & enhancement, video analysis, 3D data processing and nonlinear dimension reduction. Furthermore, we present available open source software frameworks for manifold learning. One can see from this survey that manifold learning has a lot of potential for computer vision and multimedia.

Keywords—deep learning, manifold learning, image processing, video processing

I. INTRODUCTION

Deep learning methods are nowadays the best way for the automatic analysis of multimedia data (e.g. images, video or 3D data) for tasks like classification or detection. However, classic neural networks are restricted to data lying in vector spaces, while data residing in smooth non-Euclidean spaces arise naturally in many problem domains. For example, a 360° camera actually captures a spherical image, not a rectangular image. We will focus in the following on manifolds, especially Riemannian manifolds, which are well suited for generalizing a vector space because they are locally Euclidian and differentiable.

A *manifold* M of dimension d corresponds to a topological structure which locally (so in the neighborhood of a point $p \in M$) looks like a d -dimensional Euclidean space. The "best" local approximation of this neighborhood of p with a d -dimensional Euclidean space is its *tangent space* T_pM . The tangent space T_pM can be seen as a linear approximation of M around p . For example, for a 2-dimensional manifold its tangent space T_pM is the tangent plane going through this point (see Fig. 1). A *Riemannian manifold* is a smooth manifold M equipped with a positive definite inner product g_p on the tangent space T_pM of each point p .

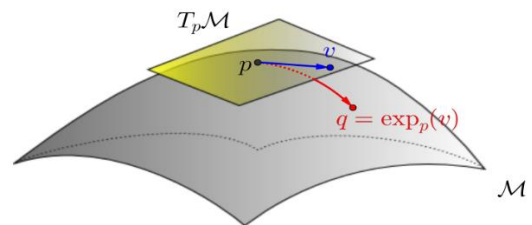


Figure 1. Tangent space and exponential map on a 2-dimensional manifold. Image courtesy of [1].

The inner product g induces a norm on the tangent space, which subsequently allows us to calculate curve lengths and distances on the manifold M . For each curve $c(t)$ on the manifold its length can be calculated by integrating the norm along the curve [2–6]. A *geodesic* curve is a *length-minimizing* curve connecting two points p and q on the manifold. The distance between these points is defined as the length of the geodesic. Let p be a (reference) point on the manifold and v a vector of its tangent space T_pM . The vector v can be mapped now to the point q on the manifold that is reached after unit time $t = 1$ by the geodesic $c(t)$ starting at p with tangent vector v . This mapping $exp_p(v) : T_pM \rightarrow M$ is called the exponential map at point p .

The inverse mapping $log_p(q) : M \rightarrow T_pM$ is uniquely defined around a neighborhood of p . Informally, the exponential map and logarithm map move points back and forth between the manifold and the tangent space (see Fig. 1) while preserving distances. Furthermore, derivative operators like *differential*, *intrinsic gradient*, *divergence* and *laplacian* can be also defined on a manifold [7, 8], which allows us to perform calculus on the manifold.

Closely related to manifolds are Lie groups. A *Lie group* is a smooth manifold that also forms a *group* [3], where both group operations (commonly called *multiplication* and *inverse*) are smooth mappings of manifolds. The *Lie algebra* \mathfrak{g} of a Lie group M is defined as the tangent space at the identity T_eM , where e is the identity element of the group (see section 16 in [8]).

Key components of neural networks – like mean, convolution, nonlinearities and batch normalization – can be defined on Riemannian manifolds as described in [9–13]. Optimization algorithms for Riemannian manifolds (gradient descent, SGD, Adam etc.) can be found in [14–21].

Commonly encountered examples of Riemannian manifolds in computer vision are the n -sphere S^n , the manifold of $n \times n$ symmetric positive matrices P_n , the special orthogonal group $SO(n)$ (rotation matrices), the special euclidean group $SE(n)$ (rigid body transformations), Grassman manifold $Gr(n,p)$ (collection of all p -dimensional linear subspaces in \mathbb{R}^n , see [22]) and the Stiefel manifold $St(n,p)$ (collection of all p -dimensional orthogonal bases in \mathbb{R}^n).

In the following, we will give an overview of manifold learning methods employed in important application fields in multimedia (similarity search, image classification, synthesis & enhancement, video analysis, 3D data processing, nonlinear dimension reduction) and about available open source software frameworks.

II. SIMILARITY SEARCH AND RETRIEVAL

Image retrieval deals with searching for similar images in an image gallery, given a certain query image (see the surveys [23, 24]). Many methods employ for this *metric learning*, which transforms input images into *embeddings* (\approx feature vectors) and learns a distance function between these embeddings.

The authors of [25] propose *regularized ensemble diffusion* for refining/reranking the initial similarity search results. They show that regularized ensemble diffusion is significantly more robust against noise in the data than standard diffusion. A *diffusion* process [26] models the relationship between objects on a graph-based manifold, wherein similarity values are diffused along the geodesic path in an iterative way.

In [27] an unsupervised framework is presented for the identification of *hard training examples* for the training of an embedding. Hard training examples (both positive and negative samples) are identified by disagreement between euclidean and manifold similarities.

A time- and memory-efficient algorithm for estimating similarities on the data manifold is proposed in [28]. They adapt the random walk procedure to estimate manifold similarities only on a small number of data in each mini-batch, rather than on all training data.

The *MLS³RDUH* algorithm [29] utilizes the intrinsic manifold structure in the feature space and cosine similarity to reconstruct the local semantic structure and build a similarity matrix upon it. Then a novel *log-cosh* hashing loss function is used to optimize the hashing network in order to generate compact hash codes, guided by the similarity matrix.

The work of [30] proposes a unsupervised metric learning algorithm that learns a metric in a lower dimensional latent space using constraints provided as tuples, which rely on pseudo-labels obtained by a graph-based clustering method (*authority ascent shift*). The parameters of the approach are learned jointly using Riemannian optimization on a product manifold.

III. IMAGE CLASSIFICATION AND OBJECT DETECTION

The work [19] proposes a framework for the transformation of problems with manifold constraints into

unconstrained problems on an Euclidean space through a mechanism they call *dynamic trivializations*. They show how to implement these trivializations efficiently for a large variety of commonly used matrix manifolds and provide a formula for the gradient of the matrix exponential.

The authors of [31] propose *manifold mixup*, a novel regularizer which forces the training to interpolate between hidden representations – captured in the intermediate layers of the network – of samples. It can be seen as a generalization of input mixup which does the interpolation on a random layer of the network (whereas input mixup uses always layer 0). Experiments for the task of image classification show that manifold mixup flattens the class-specific representation (lower variance) and generates a smoother decision boundary.

In [32] *Hyperbolic Busemann learning* with ideal prototypes is introduced. It places class prototypes at the ideal boundary of the *Poincare ball* (a hypersphere manifold with hyperbolic geometric) and introduces the *penalized Busemann loss* for optimizing with respect to ideal prototypes.

An approach for few-shot image classification is presented in [33] which proposes *embedding propagation* as an unsupervised non-parametric regularizer. Embedding propagation leverages interpolation between the extracted features of a neural network, based on a similarity graph. Experiments show that embedding propagation yields a smoother embedding manifold and gives better performance on three standard datasets for few-shot image classification.

The work [34] introduces a knowledge distillation method which is able to transfer an existing CNN model trained on perspective images to *spherical* images captured with a 360° camera *without* any additional annotation effort (see Fig. 2). They train a spherical Faster R-CNN model with this method, demonstrating that an object detector for spherical images (in equirectangular projection) can be trained without any annotations in the 360° images.

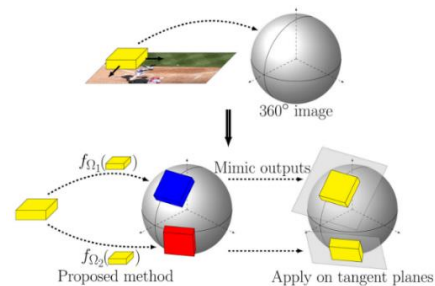


Figure 2. Transfer CNNs trained on flat images to 360° images with the method from [34].

IV. IMAGE SYNTHESIS AND ENHANCEMENT

For image synthesis and enhancement, state of the art algorithms employ either GANs (*generative adversarial networks* [35]) or *diffusion models* [36].

The authors of [37] show that current solvers employed in diffusion models throw the generative sample path off the data manifold, causing the error to

accumulate. They propose an additional correction term inspired by the manifold constraint to force the iterations to be close to the data manifold. The proposed manifold constraint is easy to add to a solver, yet boosts its performance significantly.

In [38] a novel implicit data augmentation approach for training GANs is proposed which facilitates stable strong and synthesizes high-quality samples. Specifically, the discriminator is interpreted as a metric embedding of the real data manifold, which offers real distances between real data samples. Experiments show that the proposed method improves the performance of image synthesis in the low-data regime.

A method for comparing data manifolds based on their topology is presented in [39]. They introduce novel tools, specifically *cross-barcode* and *manifold topology divergence score*, which are able to track spatial discrepancies between manifolds on multiple scales. They apply it to assess the performance of generative models in various domains (images, 3D shapes or time series) and demonstrate that these tools are able to detect common problems of GAN-based image synthesis like mode dropping, mode collapse and image disturbance.

The work [40] proposes *progressive attentional manifold alignment* for style transfer, which progressively aligns content manifolds to their most related style manifolds. Afterwards, *space-aware interpolation* is performed in order to increase the structural similarity of the corresponding manifolds, which makes it easier for the attention module to match features between them. Experiments show that the method generates high-quality style-transferred images (see Fig. 3).

The authors of [41] proposes an algorithm for improving the diversity and visual quality of images generated by a conditional GAN, by systematically encouraging a *bi-lipschitz* mapping between the latent and output manifold. The performance improvement is shown on several image-to-image translation tasks.

The FLAME algorithm proposed in [42] performs highly realistic image manipulations (e.g. changing expression, hair style or age of a synthetic face, see Fig. 4) with minimal supervision. It estimates linear latent directions in the latent space of *StyleGAN2* using only a few image pairs and introduces a novel method for sampling from the attribute style manifold.



Figure 3. From left to right: Content image, style image, style-transferred image [40].

V. VIDEO ANALYSIS

Most manifold learning methods for video analysis deal with the important task of human action recognition. Often they employ neural networks over the manifold P_n of symmetric positive matrices (usually covariance matrices) for this.

The authors of [13] propose a *dilated convolution* operator on manifolds, based on the *weighted Frechet mean* [11], as well as a *residual connection* operator. Both are important building blocks of modern neural networks. They construct a manifold-valued network employing covariance matrices (calculated from CNN features) and train this network for human action detection on the UCF-11 video dataset.

In [9] the convolution is defined as the weighted sum (reprojected to the manifold) in the tangent space T_aM , where a is the Frechet mean of the input points for the convolution. They show that their proposed convolution operator is an isometry of the manifold, which corresponds to the translation-invariance property of the convolution in an Euclidean space.

The work [43] proposes a geometry-aware deep learning algorithm for skeleton-based action recognition, where skeleton sequences are modeled as trajectories on *Kendall's shape space* and then fed into a CNN-LSTM network. Kendall's shape space is a special quotient manifold that defines shape as the geometric information that remains when location, scaling and rotational effects are filtered out [44, 45].

The algorithm [46] adopts a neural network over the manifold P_n of symmetric positive definite matrices as the backbone and appends a cascade of *Riemannian autoencoders* to it in order to enrich the information flow within the network. Experiments on the tasks of emotion recognition, hand action recognition and human action recognition demonstrate a favourable performance compared to state of the art methods.

VI. 3D DATA PROCESSING

The work [47] proposes a novel algorithm for geometric disentanglement (separate intrinsic and extrinsic geometry) of 3D models, based on the fundamental theorem for surfaces. They describe surface features via a combination of *conformal factors* and surface normal vectors and propose a convolutional mesh autoencoder based on these features. The conformal factor defines a conformal (angle-preserving) deformation between two manifolds. The algorithm achieves state-of-the-art performance on 3D surface generation, reconstruction and interpolation tasks (see Fig. 5).

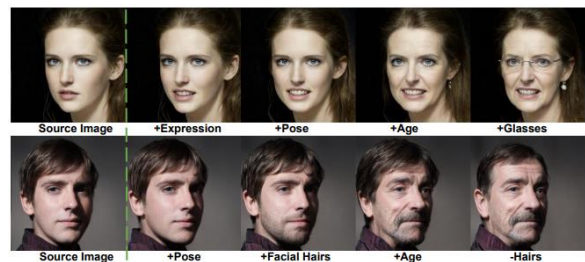


Figure 4. Image editing with FLAME [42].

The authors of [48] propose an approach for learning generative models on manifolds by minimizing the *probability path divergence*. Unlike other continuous

flow approaches, it does not require solving an ordinary differential equation during training. Furthermore, it scales well also for high dimensions.

In [49] a method for rotation (pose) estimation of 3D objects from point clouds and images is presented. For this, they propose a novel *manifold-aware* gradient in the backward pass of rotation regression that directly updates the neural network weights.

The work [50] introduces *intrinsic neural fields*, a novel and versatile representation for neural fields on manifolds. Intrinsic neural fields are based on the eigenfunctions of the *Laplace-Beltrami* operator, which can represent detailed surface information directly on the manifold. Furthermore, they extend *neural tangent kernel analysis* to manifolds for better insight into the spectral properties of neural fields.

VII. NONLINEAR DIMENSION REDUCTION

Many real world high-dimensional datasets are actually lying in a low-dimensional manifold (*manifold hypothesis*). Nonlinear dimensional reduction algorithms project high-dimensional data onto such a low-dimensional manifold, while trying to preserve distance relationships in the original high-dimensional space as good as possible.

Classical approaches for nonlinear dimension reduction are Isomap, Local Linear Embedding (LLE) and Laplacian Eigenmaps (see the survey in [51]). In recent years, more powerful approaches like *t-SNE*, *UMAP*, *TriMAP* and *PaCMAP* have emerged [52]. From these, PaCMAP seems to preserve best both the global and local structure of the high-dimensional data.

In [53], the *h-NNE* algorithm is proposed, which is competitive with t-SNE and UMAP in quality while being on order of magnitude faster. The significant runtime advantage is possible as h-NNE avoids solving an optimization problem and relies on *nearest neighbor graphs* instead.

For measuring the intrinsic dimension of a data distribution, in [54] a method is presented based on recent progress in likelihood estimation in high dimensions via *normalizing flows*.



Figure 5. Generated 3D models with the geometric disentanglement algorithm from [46].

The *SpaceMAP* algorithm introduces the concept of *equivalent extended distance*, which makes it possible to match the capacity between two spaces of different dimensionality [55] (see Fig. 6). Furthermore, *hierarchical manifold approximation* is performed based as real-world data has often a hierarchical structure.

The *DIPOLE* algorithm proposed in [56] corrects an initial embedding (e.g. calculated via Isomap) by minimizing a loss functional with both a local, metric term and a global, topological term based on *persistent homology*. Unlike more ad hoc methods for measuring the shape of data at multiple scales, persistent homology is rooted in algebraic topology and enjoys strong theoretical foundations.

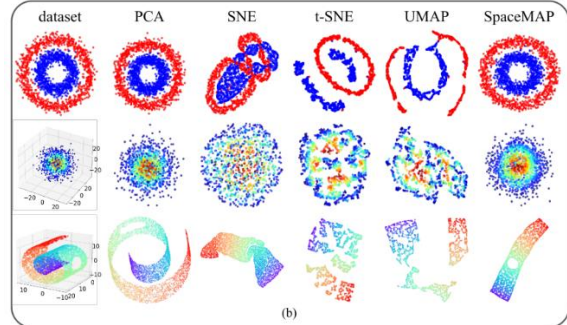


Fig. 6. Comparison of classic nonlinear dimension reduction methods with SpaceMAP [53].

VIII. OPEN SOURCE SOFTWARE PACKAGES

The Python packages *Geomstats* [57, 58], *geoopt* [16] and *Pymanopt* [59] provide implementation of the standard operators (norm, distance, exp, log, retraction, parallel transport etc.) for commonly used manifolds like S^n , P_n , $SO(n)$, $SE(n)$, $Gr(n,p)$ and $St(n,p)$.

Geomstats and *geoopt* support also more exotic manifolds like Birkhoff polytope [60], stereographic projection model, Kendall's shape space [44, 45], Poincare polydisc or hyperbolic space. Furthermore, *geoopt* provides optimizers like SGD or Adam and the sampling from a probability distribution on the manifold, whereas *Geomstats* provides Frechet mean estimators, *K*-means, and principal component analysis.

Theseus [61] provides *differentiable* optimizers (GaussNewton, Levenberg-Marquardt) and solvers (dense and sparse versions of Cholesky and LU) as well as the manifolds $SO(3)$ and $SE(3)$ which are often used in 3D data processing, robotics and kinematics. The differentiability of the optimizers/solvers makes it possible to include them into a neural network layer or loss function.

IX. CONCLUSION

Manifold learning is a novel research field in deep learning which aims to overcome the Euclidean space restriction of standard neural networks. This work provides an introduction into the key concepts of manifold learning.

An overview of important manifold learning algorithms is given for several application fields in multimedia (similarity search, image classification, synthesis & enhancement, video analysis, 3D data processing and nonlinear dimension reduction) and we furthermore describe available open source software frameworks.

CONFLICT OF INTEREST

The author declares no conflict of interest.

FUNDING

This work was supported by European Union's Horizon 2020 research and innovation programme under grant number 951911 - AI4Media.

REFERENCES

- [1] G. Miranda, C. Thomaz, and G. Giraldi, "Geometric data analysis based on manifold learning with applications for image understanding," *SIBGRAPI*, 2017.
- [2] S. Calinon, "Gaussians on riemannian manifolds: Applications for robot learning and adaptive control," *IEEE Robotics & Automation Magazine*, 2020.
- [3] T. Fletcher, "Terse notes on riemannian geometry," Technical Report, University of Utah, 2010.
- [4] S. R. Hauberg, O. Freifeld, and M. Black, "A geometric take on metric learning," *NeurIPS*, 2012.
- [5] F. Porikli, "Learning on manifolds," *Structural, Syntactic, and Statistical Pattern Recognition*, 2010.
- [6] S. Sommer and T. Fletcher, *Riemannian Geometric Statistics in Medical Image, Chapter Introduction to Differential and Riemannian Geometry*, Elsevier, 2020.
- [7] M. Bronstein, J. Bruna, and Y. LeCun, "Geometric deep learning: Going beyond euclidean data," *IEEE Signal Processing Magazine*, 2017.
- [8] L. W. Tu, *An Introduction to Manifolds*, Springer, 2007.
- [9] J. J. Bouza, C.-H. Yang, D. E. Vaillancourt, and B. C. Vemuri, "Mvc-net: A convolutional neural network architecture for manifold-valued images with applications," ArXiv, 2020.
- [10] R. Chakraborty, "Manifoldnorm: Extending normalizations on riemannian manifolds," ArXiv, 2020.
- [11] R. Chakraborty, J. Bouza, J. H. Manton, and B. C. Vemuri, "Manifoldnet: A deep neural network for manifold-valued data with applications," *IEEE TPAMI*, 2022.
- [12] A. Lou, I. Katsman, Q. Jiang, and S. Belongie, "Differentiating through the frechet mean," *ICML*, 2020.
- [13] X. Zhen, R. Chakraborty, and N. Vogt, "Dilated convolutional neural networks for sequential manifold-valued data," *ICCV*, 2019.
- [14] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, 2007.
- [15] F. Alimisis and A. Orvieto, "Momentum improves optimization on riemannian manifolds," *AISTATS*, 2021.
- [16] G. Becigneul and O.-E. Ganea, "Riemannian adaptive optimization methods," *ICLR*, 2019.
- [17] M. Cho and J. Lee, "Riemannian approach to batch normalization," *NeurIPS*, 2017.
- [18] H. Kasai, P. Jawanpuria, and B. Mishra, "Riemannian adaptive stochastic gradient algorithms on matrix manifolds," *ICML*, 2019.
- [19] M. L. Casado, "Trivializations for gradient-based optimization on manifolds," *NeurIPS*, 2019.
- [20] J. Li, F. Li, and S. Todorovic, "Efficient riemannian optimization on the stiefel manifold via the cayley transform," *AISTATS*, 2021.
- [21] A. Sim, M. L. Wiatrak, and A. Brayne, "Directed graph embeddings in pseudo-riemannian manifolds," *ICML*, 2021.
- [22] T. Bendokat, R. Zimmermann, and P. Absil, "A grassmann manifold handbook: Basic geometry and computational aspects," ArXiv, 2020.
- [23] W. Chen, Y. Liu, and W. Wang, "Deep learning for instance retrieval: A survey," *IEEE TPAMI*, 2022.
- [24] S. R. Dubey, "A decade survey of content based image retrieval using deep learning," *IEEE Transactions on Circuits and Systems for Video Technology*, 2022.
- [25] S. Bai, Z. Zhou, and J. Wang, "Ensemble diffusion for retrieval," *ICCV*, 2017.
- [26] M. Donoser and H. Bischof, "Diffusion processes for retrieval revisited," *CVPR*, 2013.
- [27] A. Iscen, G. Tolas, and Y. Avrithis, "Mining on manifolds: Metric learning without labels," *CVPR*, 2018.
- [28] N. Aziere and S. Todorovic, "Ensemble deep manifold similarity learning using hard proxies," *CVPR*, 2019.
- [29] R.-C. Tu, X.-L. Mao, and W. Wei, "Mls3rduh: Deep unsupervised hashing via manifold based local semantic similarity structure reconstructing," *IJCAI*, 2020.
- [30] U. K. Dutta and C. Sekha, "A geometric approach for unsupervised similarity learning," *ICASSP*, 2020.
- [31] V. Verma, A. Lamb, and Y. Bengio, "Manifold mixup: Better representations by interpolating hidden states," *ICML*, 2019.
- [32] M. G. Atigh, M. Keller-Ressel, and P. Mettes, "Hyperbolic busemann learning with ideal prototypes," *NeurIPS*, 2021.
- [33] P. Rodriguez, I. Laradji, and A. Drouin, "Embedding propagation: Smoother manifold for few-shot classification," *ECCV*, 2020.
- [34] Y.-C. Su and K. Grauman, "Learning spherical convolution for 360° recognition," *IEEE TPAMI*, 2022.
- [35] Z. Wang, Q. She, and T. Ward, "Generative adversarial networks in computer vision," *ACM Computing Surveys*, 2021.
- [36] F.-A. Croitoru and V. Hondru, "Diffusion models in vision: A survey," ArXiv, 2022.
- [37] H. Chung, B. Sim, and J. C. Ye, "Improving diffusion models for inverse problems using manifold constraints," 2022.
- [38] M. Dai, H. Hang, and X. Guo, "Adaptive feature interpolation for low-shot image generation," *ECCV*, 2022.
- [39] S. Barannikov, I. Trofimov, and G. Sotnikov, "Manifold topology divergence: a framework for comparing data manifolds," *NeurIPS*, 2021.
- [40] X. Luo, Z. Han, and L. Yang, "Progressive attentional manifold alignment for arbitrary style transfer," *ACCV*, 2022.
- [41] S. Ramasinghe and M. R. Farazi, "Rethinking conditional GAN training: An approach using geometrically structured latent manifolds," *NeurIPS*, 2021.
- [42] R. Parihar, A. Dhiman, and T. Karmali, "Everything is there in latent space: Attribute editing and attribute style manipulation by stylegan latent space exploration," *ACM MM*, 2022.
- [43] R. Frijji, H. Drira, and F. Chaieb, "KShapeNet: Riemannian network on kendall shape space for skeleton based action recognition," *CoRR*, 2020.
- [44] N. Guigui, E. Maignant, A. Trouve, and X. Pennec, "Parallel transport on kendall shape spaces," *GSI*, 2021.
- [45] D. Kendall, D. M. Barden, and T. K. Carne, "Shape and shape theory," 1999.
- [46] R. Wang, X.-J. Wu, Z. Chen, and T. Xu, "Dreamnet: A deep riemannian manifold network for spd matrix learning," *ACCV*, 2022.
- [47] S. C. N. J. Tatro, "Unsupervised geometric disentanglement for surfaces via cfan-vae," *ICLR 2021 Workshop on Geometrical and Topological Representation Learning*, 2021.
- [48] H. Ben-Hamu, S. Cohen, and J. Bose, "Matching normalizing flows and probability paths on manifolds," *ICML*, 2022.
- [49] J. Chen, Y. Yin, and T. Birdal, "Projective manifold gradient layer for deep rotation regression," *CVPR*, 2022.
- [50] L. Koestler *et al.*, "Intrinsic neural fields: Learning functions on manifolds," *ECCV*, 2022.
- [51] L. Cayton, "Algorithms for manifold learning," *CoRR*, 2005.
- [52] Y. Wang, H. Huang, C. Rudin, and Y. Shaposhnik, "Understanding how dimension reduction tools work: An empirical approach to deciphering t-SNE, UMAP, TriMap, and PaCMAP for data visualization," *J. Mach. Learn. Res.*, 2022.
- [53] M. S. Sarfraz, M. Koulakis, C. Seibold, and R. Stiefelhagen, "Hierarchical nearest neighbor graph embedding for efficient dimensionality reduction," *CVPR*, 2022.
- [54] P. Tempczyk, R. Michaluk, and L. Garncarek, "LIDL: Local intrinsic dimension estimation using approximate likelihood," *ICML*, 2022.
- [55] X. Zu and Q. Tao, "SpaceMAP: Visualizing high-dimensional data by space expansion," *ICML*, 2022.
- [56] A. Wagner, E. Solomon, and P. Bendich, "Improving metric dimensionality reduction with distributed topology," ArXiv, abs/2106.07613, 2021.
- [57] N. Miolane, N. Guigui, A. Le Brigant, J. Mathe, B. Hou, Y. Thanwerdas, and S. Heyder, "Geomstats: A python package for riemannian geometry in machine learning," *J. Mach. Learn. Res.*, vol. 21. No. 1, 2022.
- [58] N. Miolane, N. Guigui, and H. Zaatiti, "Introduction to geometric learning in python with geomstats," *SciPy 2020*, 2020.

- [59] J. Townsend, N. Koep, and S. Weichwald, "Pymanopt: A python toolbox for optimization on manifolds using automatic differentiation," *J. Mach. Learn. Res.*, 2016.
- [60] A. Douik and B. Hassibi, "Manifold optimization over the set of doubly stochastic matrices: A second-order geometry," Technical Report, 2018.
- [61] L. Pineda, T. Fan, M. Monge, S. Venkataraman, and P. Sodhi, "Theseus: A library for differentiable nonlinear optimization," *NeurIPS*, 2022.

Copyright © 2023 by the authors. This is an open access article distributed under the terms of Creative Commons attribution-noncommercial license ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.