Deep and Convex Shape Analysis

HAGGAI MARON, Weizmann Institute of Science

In this paper, I review the main results obtained during my Ph.D. studies at the Weizmann Institute of Science under the guidance of Professor Yaron Lipman. Two fundamental problems in shape analysis were considered: (1) how to apply deep learning techniques to geometric objects and (2) how to compute meaningful maps between shapes. My work has resulted in several novel methods for applying deep learning to surfaces, point clouds, and hyper-graphs as well as new efficient techniques to solve relaxations of well-known matching problems. The paper discusses these two problems, surveys the suggested solutions and points out several directions for future work, including a promising direction that combines both problems.

Additional Key Words and Phrases: Shape Analysis, Deep Learning, Convex Relaxation, Shape Matching

ACM Reference Format:
Haggai Maron. 2019. Deep and Convex Shape Analysis. 1, 1 (July 2019), 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Shape analysis is concerned with studying and quantifying geometric properties of objects such as curves, surfaces, and higher dimensional manifolds. Among other fields, shape analysis techniques are widely used in computer vision [Zhou et al. 2016], computer graphics [Funkhouser and Kazhdan 2004], computational anatomy [Boyer et al. 2011] and medical imaging [McInerney and Terzopoulos 1996].

During the last few years we tried to tackle two key questions in shape analysis: (1) deep learning on geometric objects: *i.e.*, How can we apply deep learning to common shape representations such as point clouds, surfaces, and graphs? (2) Shape matching: Given two shapes, how can we compute a meaningful map between them. The rest of the introduction put these two problems in the proper context.

Applying deep learning to geometric objects is a relatively new problem. The overwhelming success of deep learning in advancing the state of the art in various learning challenges and domains [LeCun et al. 2015] inspires research efforts attempting to achieve similar success for geometric objects such as point-clouds, graphs and discretized surfaces [Bronstein et al. 2017]. Adapting deep learning methods to the geometric setting is a particularly interesting and challenging problem since each of these objects admits different representations, and consequently, different symmetries: for example, surfaces and point clouds (*i.e.*, finite subsets of the Euclidean space) are invariant to rigid Euclidean transformations, while graphs are invariant to node relabeling. Trying to directly apply commonly used neural networks (*e.g.*, convolutional neural networks or fully connected networks) to geometric objects is not well-defined in some cases, or performs poorly in other cases. During the last years, we developed network architectures and layers for all of the geometric objects mentioned above, as well as analyzed widely-used models.

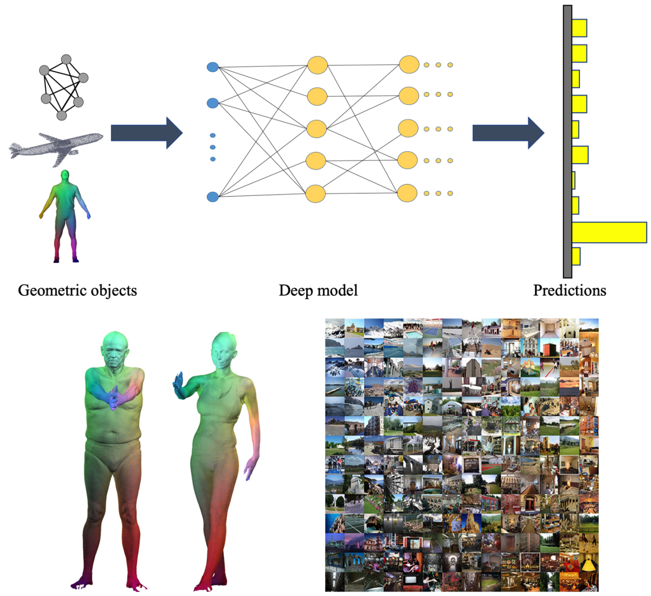

Fig. 1. The main problems considered in my Ph.D. thesis: applying deep learning to geometric objects such as point clouds, surfaces, and graphs. Bottom: Two instances of shape matching problems.

In contrast to the first problem, shape matching problems have been studied for decades [Van Kaick et al. 2011]. These problems are among the most fundamental problems in geometric data analysis, where the task is finding a (semantically) meaningful map between one shape to another shape. A popular way of handling these hard problems is by first posing them as quadratic integer optimization problems, relaxing them to a continuous domain and solving the relaxed problem, which is often more tractable. In most cases, there is an inherent tradeoff between the tightness of the relaxation (*i.e.*, how well its solution approximates the original problem's solution), and the computational resources needed to optimize it. During the last few years, we devised several efficient methods for solving widely known relaxations of prevalent matching problems.

Figure 1 illustrates these two problems. The top part illustrates the geometric deep learning problem: applying deep learning to irregular domains. The bottom part illustrates two instances of the shape matching problem: matching 3D models (left, the map is represented using color coding) and matching an image collection to a grid structure according to color features (right).

Author's address: Haggai Maron, haggaimaron@gmail.com, Weizmann Institute of Science, Rehovot, Israel.

XXXX-XXXX/2019/7-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

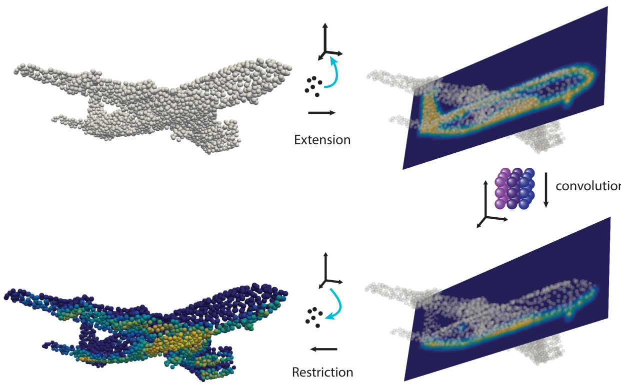

Fig. 2. The point cloud convolution suggested in [Atzmon et al. 2019]: First, a function over the point cloud (in this case the constant one) is *extended* to a continuous volumetric function over the ambient space; second, a continuous volumetric *convolution* is applied to this function (without any discretization or approximation); and lastly, the result is *restricted* back to the point cloud.

The rest of this paper is organized in the following way: Section 2 summarizes several papers that targeted the first problem. Section 3 summarizes efforts that were done in order to solve the second problem. Section 4 concludes that paper and discusses possible future work directions.

2 DEEP LEARNING ON GEOMETRIC OBJECTS

The problem setting is as follows: the input consists of n geometric objects O_i, possibly with corresponding descriptor functions $f_i : O_i \to \mathbb{R}^d$, and targets t_i that describe some semantic property of the object (*e.g.*, does a surface represent a dog or a cat). Our task is to find a function $\mathcal{F}$ that approximates this functional relation, *i.e.*, $\mathcal{F}(O_i, f_i) \approx t_i$ and generalizes well to unseen objects.

Discretized surfaces and point clouds. We first studied deep learning of discretized surfaces [Maron et al. 2017] and point clouds [Atzmon et al. 2018]. In both cases, our methods are based on the observation that finding a well-behaved mapping from the given object (surface or point cloud) to a domain with a well-defined convolution, allows us to *pullback* this convolution to the object. The pullback operation is done by first mapping a function from the object to this domain, applying the convolution and mapping the result back to the object. In [Maron et al. 2017] we devised a way to map sphere-type surfaces (*i.e.*, surfaces that are topologically equivalent to a sphere) to a periodic planar domain (a torus), for which we have the well-known 2D convolution. Using this mapping mechanism, we convert input surface descriptors to images, thus reducing the problem of learning surfaces to the problem of learning images. This reduction has the advantage of being able to utilize powerful deep learning machinery and successful architectures developed for images, as opposed to methods that are specifically tailored for surfaces (*e.g.*, [Masci et al. 2015]). In a follow-up work [Ben-Hamu et al. 2018], we used a similar surface-image representation for the task of generating novel shapes using generative adversarial networks. In particular, we

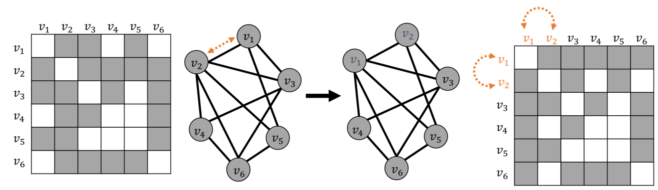

Fig. 3. An illustration of symmetries of graph representations. When representing a graph using an adjacency matrix, each ordering of the nodes gives rise to a new, possibly different adjacency matrix of the same graph. In [Maron et al. 2019a,b] we suggest neural networks that are invariant to such symmetries.

have successfully applied our method to the task of human model generation. In another recent follow-up work [Haim et al. 2018], we devise a broad family of mappings from sphere-type surfaces to the torus, which represents surfaces more faithfully due to reduced area distortion. This method achieves state of the art results on 3D shape analysis tasks such as model segmentation and retrieval.

The mapping methods mentioned above rely heavily on the surface's connectivity information and cannot be adapted to the point cloud scenario [Atzmon et al. 2019]. In this case, which is of particular interest for applications, we opt for mapping point cloud functions to functions defined on $\mathbb{R}^3$. This is done by defining an *extension* operator that generates a volumetric function from a point cloud function via a Radial Basis Function (RBF) approximation. Similarly, the kernel is defined to be a sum of weighted RBFs, where the weights are the learnable parameters. The convolution of a point cloud function can now be defined using the pullback mechanism mentioned above: first mapping the point cloud function to a volumetric function, applying the standard convolution in $\mathbb{R}^3$ and sampling the result on the point cloud in order to get new point cloud function. This process is illustrated in Figure 2.

Graphs and hyper-graphs. In a related line of work, we study a popular model for constructing networks that are invariant to natural transformations of the input object [Hartford et al. 2018; Maron et al. 2019b; Ravanbakhsh et al. 2017; Zaheer et al. 2017]. Given a group $\mathcal{G}$ acting on an input object, this model is composed of a concatenation of equivariant/invariant linear layers with nonlinearities. A fundamental problem when constructing such models is finding the maximal set of linear invariant or equivariant operators with respect to the relevant group. In [Maron et al. 2019b], we addressed this problem for the natural symmetry groups of graphs and hyper-graphs. In this case, the input is an affinity tensor $A \in \mathbb{R}^{n^k}$ that describes relations between ordered subsets of k elements in a set (*e.g.*, in the case $k = 2$, A is a graph affinity matrix). Note that these tensors adhere to a specific reordering symmetry: reordering the nodes of the hyper-graph results in a different affinity tensor that represents the same hyper-graph. Figure 3 illustrates this symmetry. In this paper, we provided a full characterization of affine functions that are equivariant to this reordering operation. One surprising fact is that the dimension of the space of equivariant functions does not depend on n, the number of nodes. This fact allowed us to construct deep invariant networks that can process graphs of any size. Theoretically,

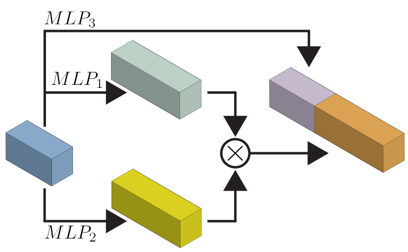

Fig. 4. The permutation equivariant block suggested in [Maron et al. 2019a]. Three different Multi-Layer Perceptrons are applied to the feature dimension of the input tensor. Matrix multiplication is applied to the output of two of them while the third output is concatenated.

we show that this construction gives rise to a deep invariant model that can approximate any *message passing neural network* [Gilmer et al. 2017], the current state of the art in graph neural networks.

In [Maron et al. 2019c], we study the approximation power of the invariant models mentioned above. We consider the rather general case of permutation groups acting on $\mathbb{R}^n$ by permuting the coordinates of vectors and show three main results. The first result, states that this model can approximate any continuous invariant function to arbitrary precision. The proof is constructive and makes use of high order tensors, that is, tensors of the form $A \in \mathbb{R}^{n^k}$ for $k > 1$, which might be computationally prohibitive. Our second result shows that this problem cannot be alleviated since there exist an infinite family of permutation groups for which using high order tensors is necessary for obtaining the universal approximation property. Our last result considers the most important case for applications, *i.e.*, networks that use only first order tensors (*e.g.*, $k = 1$), and provides a necessary condition for a permutation group to have the universal approximation property in this case.

In our latest paper [Maron et al. 2019a] we give more evidence to the fact that higher-order networks are more expressive, and suggest to deviate from the linear equivariant model that was suggested above for achieving better complexity vs. expressivity tradeoff. More precisely, we show that a k-order networks can approximate the k-Weisfeiler Lehman (k-WL) graph isomorphism test [Xu et al. 2018], which gives rise to graph models that are more powerful than message passing neural networks. We also suggest a new simple architecture, composed of blocks that apply Multi Layer Perceptrons (MLP) to the edge and node features and then matrix multiplication. This model is shown to have 3-WL expressivity, strictly more powerful than message passing networks. See Figure 4 for an illustration of the suggested block.

3 RELAXATIONS OF MATCHING PROBLEMS

Three of our works devise scalable approaches for solving well-known tight relaxations of popular matching problems. In all cases, we started from a classic semidefinite relaxation [Helmberg 2000] in which the quadratic terms are linearized at the cost of adding a new large optimization variable. In general, this is a tight relaxation that can be solved efficiently for only small-sized problems.

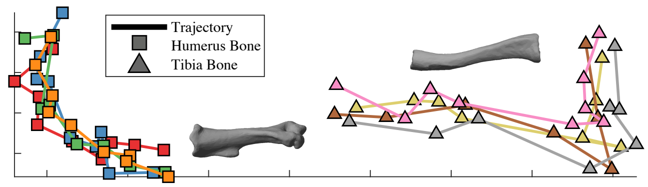

Fig. 5. Anatomical dataset embedding in the plane obtained using the method suggested in [Maron and Lipman 2018]. Squares and triangles represent different bone types; lines represent temporal trajectories.

In [Maron et al. 2016] the problem of jointly aligning and matching two point clouds is considered. More precisely, Given two d-dimensional point clouds, $P, Q \in \mathbb{R}^{d\times n}$, which are neither aligned nor consistently labeled, the task is to find an orthogonal transformation $R \in O(d)$ and a permutation $X \in \Pi_n$ minimizing the distance between the point clouds:

$$\begin{aligned} d(P,Q) = \min_{X,R} & \|RP - QX\|_F^2 \\ \text{s.t.} \quad & X \in \Pi_n \\ & R \in O(d) \end{aligned}$$

This is a central problem in shape analysis with many applications in computer vision and computer graphics. Applying the standard semidefinite relaxation results in a relaxation that can be solved for up to 15 points. Our key insight is that the large semidefinite constraint can be shown to be equivalent to several smaller semidefinite constraints. This observation allowed us to solve problems with significantly larger number of points.

Our motivation in [Dym et al. 2017] was to find an efficient way of optimizing a tight relaxation to the graph matching problem (GM). Given two graphs represented by affinity matrices $A, B \in \mathbb{R}^{n\times n}$ the GM problem is the problem of finding a permutation matrix X that minimizes a quadratic objective that measures the discrepancy between edge affinities of the graph, *e.g.*,

$$E(X) = -\mathrm{tr}(AXBX^T)$$

As in [Maron et al. 2016], applying the standard semidefinite relaxation is impractical for real-sized problems. The key contribution of this paper is showing that this semidefinite relaxation is equivalent to a convex quadratic program, which can be solved more efficiently. Using this observation, we optimize this relaxation for graphs of much larger size. Figure 1 (bottom) shows results that were obtained by this method. In [Kushinsky et al. 2019], we show how to approximately minimize another relaxation that originates from the lifting method, by using a Sinkhorn-type method [Cuturi 2013].

In [Maron and Lipman 2018], we analyze the common doubly-stochastic relaxation for the GM problem. In this case, the domain of the relaxation is the convex hull of all permutation matrices, and the objective is again $E(X)$. Our first result shows that many instances of this relaxation, *e.g.*, when matching graphs represented by Euclidean distance affinities, are concave relaxations. This is a significant result since concave relaxations have two important advantages: (i) every local minimum is a permutation matrix and (ii) the set of global minima of the original problem and the relaxation is the same. Differently put, the relaxation process does not yield new solutions as well as alleviates the need to project the solution of the relaxed problem onto the permutation matrices (a step which is often not optimal). Our second result shows that many other popular use cases, *e.g.*, when matching graphs represented by geodesic distance affinity matrices, are concave with high probability, meaning it is rare to find a direction on which the restriction of the objective is convex. we also show that in these cases the relaxation enjoys the advantages mentioned above with high probability. Figure 5 illustrates an application of this concave relaxation to anatomical shape space analysis: We match a dataset of 67 mice bone surfaces acquired using micro-CT. The dataset consists of eight time series. Each time series captures the development of one type of bone over time. We used Multi-Dimensional Scaling (MDS) [Cox and Cox 2000] to assign 2D coordinates to each surface using a dissimilarity matrix we obtained from matching all pairs of bones.

4 CONCLUSION

Although considerable progress was obtained in the last few years, both problems considered in this paper are far from being solved. As for the problem of deep learning on geometric objects, there are no methods that can work on all types of meshes, including triangle soups which are abundant in applications. A possible way to tackle this problem is using the hyper-graph learning approach from [Maron et al. 2019b] as was recently suggested by [Albooyeh et al. 2019]. As for matching problems, there is still no silver bullet solution for matching three-dimensional shapes. One prominent direction is to learn how to efficiently solve these hard optimization problems, as was suggested in *e.g.*[Li et al. 2018; Masci et al. 2015; Wei et al. 2016].

Acknowledgements

The author would like to thank Professor Yaron Lipman for his guidance, constant encouragement, and advice. I was fortunate to have him as an advisor. The work discussed in this paper was supported in part by the European Research Council (ERC Starting Grant "Surf-Comp" 307754, ERC Consolidator Grant "LiftMatch" 771136), the Israel Science Foundation (Grant No. 1284/12,1830/17) and I-CORE program of the Israel PBC and ISF (Grant No. 4/11).

REFERENCES

Marjan Albooyeh, Daniele Bertolini, and Siamak Ravanbakhsh. 2019. Incidence Networks for Geometric Deep Learning. *arXiv preprint arXiv:1905.11460* (2019).

Matan Atzmon, Niv Haim, Lior Yariv, Ofer Israelov, Haggai Maron, and Yaron Lipman. 2019. Controlling Neural Level Sets. *arXiv preprint arXiv:1905.11911* (2019).

Matan Atzmon, Haggai Maron, and Yaron Lipman. 2018. Point Convolutional Neural Networks by Extension Operators. *ACM Trans. Graph.* 37, 4, Article 71 (July 2018), 12 pages.

Heli Ben-Hamu, Haggai Maron, Itay Kezurer, Gal Avineri, and Yaron Lipman. 2018. Multi-chart Generative Surface Modeling. *ACM Trans. Graph.* 37, 6, Article 215 (Dec. 2018), 15 pages.

Doug M Boyer, Yaron Lipman, Elizabeth St Clair, Jesus Puente, Biren A Patel, Thomas Funkhouser, Jukka Jernvall, and Ingrid Daubechies. 2011. Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proceedings of the National Academy of Sciences* 108, 45 (2011), 18221–18226.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. 2017. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine* 34, 4 (2017), 18–42.

Trevor F Cox and Michael AA Cox. 2000. *Multidimensional scaling*. Chapman and hall/CRC.

Marco Cuturi. 2013. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in neural information processing systems*. 2292–2300.

Nadav Dym, Haggai Maron, and Yaron Lipman. 2017. DS++: A Flexible, Scalable and Provably Tight Relaxation for Matching Problems. *ACM Trans. Graph.* 36, 6, Article 184 (Nov. 2017), 14 pages.

Thomas Funkhouser and Michael Kazhdan. 2004. Shape-based retrieval and analysis of 3D models. In *ACM SIGGRAPH 2004 Course Notes*. ACM, 16.

Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. 2017. Neural message passing for quantum chemistry. *arXiv preprint arXiv:1704.01212* (2017).

Niv Haim, Nimrod Segol, Heli Ben-Hamu, Haggai Maron, and Yaron Lipman. 2018. Surface Networks via General Covers. *arXiv preprint arXiv:1812.10705* (2018).

Jason Hartford, Devon R Graham, Kevin Leyton-Brown, and Siamak Ravanbakhsh. 2018. Deep Models of Interactions Across Sets. *arXiv preprint arXiv:1803.02879* (2018).

Christoph Helmberg. 2000. Semidefinite programming for combinatorial optimization. (2000).

Yam Kushinsky, Haggai Maron, Nadav Dym, and Yaron Lipman. 2019. Sinkhorn Algorithm for Lifted Assignment Problems. *SIAM Journal on Imaging Sciences* 12, 2 (2019), 716–735.

Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436.

Zhuwen Li, Qifeng Chen, and Vladlen Koltun. 2018. Combinatorial optimization with graph convolutional networks and guided tree search. In *Advances in Neural Information Processing Systems*. 537–546.

Haggai Maron, Heli Ben-Hamu, Hadar Serviansky, and Yaron Lipman. 2019a. Provably Powerful Graph Networks. *arXiv preprint arXiv:1905.11136* (2019).

Haggai Maron, Heli Ben-Hamu, Nadav Shamir, and Yaron Lipman. 2019b. Invariant and Equivariant Graph Networks. In *International Conference on Learning Representations*. https://openreview.net/forum?id=Syx72jC9tm

Haggai Maron, Nadav Dym, Itay Kezurer, Shahar Kovalsky, and Yaron Lipman. 2016. Point registration via efficient convex relaxation. *ACM Transactions on Graphics (TOG)* 35, 4 (2016), 73.

Haggai Maron, Ethan Fetaya, Nimrod Segol, and Yaron Lipman. 2019c. *On the Universality of Invariant Neural Networks*. Technical Report.

Haggai Maron, Meirav Galun, Noam Aigerman, Miri Trope, Nadav Dym, Ersin Yumer, Vladimir G Kim, and Yaron Lipman. 2017. Convolutional neural networks on surfaces via seamless toric covers. *ACM Trans. Graph* 36, 4 (2017), 71.

Haggai Maron and Yaron Lipman. 2018. (Probably) Concave Graph Matching. In *Advances in Neural Information Processing Systems*. 406–416.

Jonathan Masci, Davide Boscaini, Michael Bronstein, and Pierre Vandergheynst. 2015. Geodesic convolutional neural networks on riemannian manifolds. In *Proceedings of the IEEE international conference on computer vision workshops*. 37–45.

Tim McInerney and Demetri Terzopoulos. 1996. Deformable models in medical image analysis: a survey. *Medical image analysis* 1, 2 (1996), 91–108.

Siamak Ravanbakhsh, Jeff Schneider, and Barnabas Poczos. 2017. Equivariance through parameter-sharing. *arXiv preprint arXiv:1702.08389* (2017).

Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. 2011. A survey on shape correspondence. In *Computer Graphics Forum*, Vol. 30. Wiley Online Library, 1681–1707.

Lingyu Wei, Qixing Huang, Duygu Ceylan, Etienne Vouga, and Hao Li. 2016. Dense human body correspondences using convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1544–1553.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. 2018. How Powerful are Graph Neural Networks? *arXiv preprint arXiv:1810.00826* (2018).

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan R Salakhutdinov, and Alexander J Smola. 2017. Deep sets. In *Advances in Neural Information Processing Systems*. 3391–3401.

Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. 2016. Fast global registration. In *European Conference on Computer Vision*. Springer, 766–782.