



HAL
open science

Characterizing Images by the Gromov-Hausdorff Distances Between Derived Hierarchies

Amin Fehri, Santiago Velasco-Forero, Fernand Meyer

► **To cite this version:**

Amin Fehri, Santiago Velasco-Forero, Fernand Meyer. Characterizing Images by the Gromov-Hausdorff Distances Between Derived Hierarchies. International Conference on Image Processing (ICIP), Oct 2018, Athens, Greece. hal-01823866

HAL Id: hal-01823866

<https://hal.science/hal-01823866v1>

Submitted on 9 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CHARACTERIZING IMAGES BY THE GROMOV-HAUSDORFF DISTANCES BETWEEN DERIVED HIERARCHIES

Amin Fehri, Santiago Velasco-Forero and Fernand Meyer

MINES ParisTech, PSL Research University, CMM - Center of Mathematical Morphology
 {amin.fehri, santiago.velasco, fernand.meyer@mines-paristech.fr}

ABSTRACT

A hierarchy is a series of nested partitions in which a coarser partition results from merging regions of finer ones. Each hierarchy derived from an image provides a particular structural description of the image content, depending upon the criteria for merging neighboring regions. Distinct hierarchies derived from a same image reflect its various facets and the distances between them nicely characterize its content. In this paper the hierarchies are constructed with the versatile stochastic watershed algorithm and their inter-distances are measured with the Gromov-Hausdorff distance. Experiments conducted on images simulated by dead leaves model illustrate the advantages of our approach in terms of learning efficiency and understandability of the results.

Index Terms— Gromov-Hausdorff Distance, Hierarchies, Stochastic Watershed, Texture, Classification

1. INTRODUCTION

Hierarchical segmentation is a multi-scale image analysis technique that provides contours valuations proposals, with respect to given properties to enhance. It has been one of the fruitful approaches to perform segmentation tasks. The goal then is to find the hierarchical segmentation from which it is easier to find a suitable segmentation [1, 2, 3, 4, 5]. However, these different hierarchical segmentation approaches are application-dependent, and their efficiency depends on the types of structures present in the images. In this paper, we propose a new approach to take advantage of such multi-scale analysis. The main idea is to quantify the distance between different hierarchical representations to characterize images. In the proposed approach, we only use selected morphological hierarchical segmentation methods [6]. However, any hierarchical method can fall within its scope.

First, we remind the reader of ways to generate hierarchical segmentations, which basically consist in progressive merging of images regions such that the regions we are interested in are the last to merge with their neighbors. Then starting from a set of hierarchies and a given image, we show how we can compute a distance between them: we use this

representation to characterize classes of images in an analytical way. Finally, we illustrate the discriminative and explanatory power of these features on generated shape images. An overview of the process can be found in Fig. 1. Our main contributions are: (i) The use of the Gromov-Hausdorff distance to quantify the relative contributions of each hierarchy of segmentation, (ii) The introduction of image features in the form of interhierarchy distance matrices.

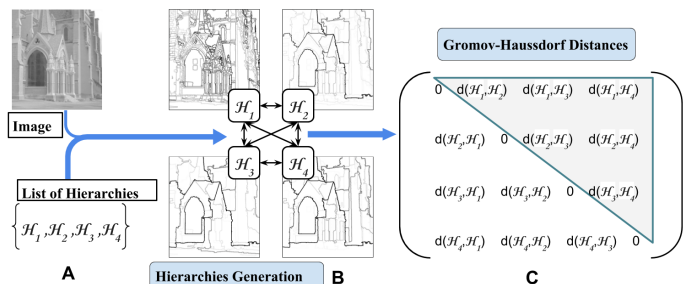


Fig. 1. Overview of our approach **A**: We start from an image and a set of morphological hierarchical segmentations techniques. **B**: Each of these techniques highlight different contours of the image, and we can quantify their relative contribution using the Gromov-Hausdorff distance. **C**: We proposed to use the interhierarchy Gromov-Hausdorff distance matrices as features.

2. MORPHOLOGICAL HIERARCHIES

2.1. Hierarchies of Segmentations

Hierarchical segmentation is a low-level image analysis often used to make easier the obtention of a suitable segmentation. In this work, we limit ourselves to morphological hierarchies that can efficiently be obtained in a graph-based framework [7]. For the sake of brevity, the reader is invited to refer to the section 2. of [8] for a description of this process. Making use of such approaches, we have several ways to generate an *indexed hierarchy of partitions* (\mathcal{H}, λ) , with \mathcal{H} a chain of nested partitions $\mathcal{H} = \{\pi_0, \pi_1, \dots, \pi_n | \forall j, k, 0 \leq j \leq k \leq n \Rightarrow \pi_j \sqsubseteq \pi_k\}$, with π_n the single-region partition, π_0

the finest partition on the image and $\pi_i \sqsubseteq \pi_j$ if all regions of π_i are included in regions of π_j , and $\lambda : \mathcal{H} \rightarrow \mathbb{R}^+$ being an ultrametric stratification index verifying $\lambda(\pi) < \lambda(\pi')$ for two nested partitions $\pi \subset \pi'$. This increasing map allows us to value each contour according to the level of the hierarchy for which it disappears: this is the *saliency* of the contour [5]. A contour then has various saliencies for various hierarchies. An example of this is shown in Fig. 1B.

2.2. A wide range of hierarchies

Based on the framework described in the previous section, one can generate several types of hierarchies that highlight different contours in the image. In particular, the model of the stochastic watershed (SWS) hierarchies [6], based upon the seminal work of [9] is able to adapt to the specific object properties to enhance (e.g. elongation, surfaces equilibrium, contrast) for segmentation tasks [8]. It can also take spatial prior information into account in its making [10]. If we consider our image as a topographic relief, flooding this image leads to watershed lines, i.e. to a segmentation. By spreading random flooding sources multiple times and flooding the image accordingly, one can characterize each contour of the image by its frequency of appearance in the associated segmentations. By only choosing high-level parameters such as the density governing the distribution of markers or their shape and size, we can then generate at will various multilevel representations of the images highlighting various types of regions. Thanks to this versatility, specific hierarchies can be built: for example, we can favor certain shapes in the images by considering a surface-based SWS hierarchy in which the measured areas are the ones of regions eroded with anisotropic structuring elements. An example of this is shown in Fig. 2: vertical structures are highlighted in (c), and horizontal ones in (d). An additional layer of supplementary complexity can be added via the possibility to combine hierarchies to obtain new ones. It can notably be done through the use of subdominant ultrametries to obtain the supremum, infimum or a linear combination of two hierarchies [11, 7]. In this work, when saliencies are probability values (as it is the case of stochastic hierarchies), the same procedure can lead to probabilistic combinations of hierarchies of the type $\text{AND}(\mathcal{H}_1, \mathcal{H}_2)$ and $\text{OR}(\mathcal{H}_1, \mathcal{H}_2)$. In Fig. 2(f), the circular and cross shapes are highlighted, as they respect the horizontal and vertical property. We argue that this wealth of controlled understandable options can lead to powerful features, through the use of a distance between hierarchies: the Gromov-Hausdorff distance.

3. FEATURES ON HIERARCHIES USING GROMOV-HAUSDORFF DISTANCE

Indeed, once provided with a family of hierarchies, one may wonder if there are ways to use the different information they provide to characterize images. The usual approach to do so

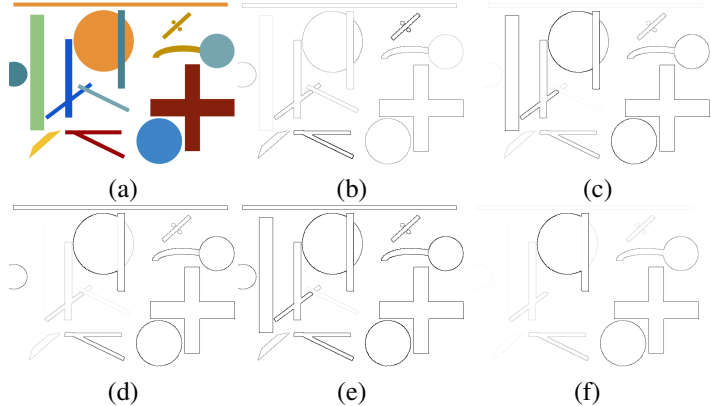


Fig. 2. Saliency maps for different hierarchies (inverted for better visualization). (a) Image, (b) Dynamics, (c) $\mathcal{H}_{surf-VertSE}$, (d) $\mathcal{H}_{surf-HorizSE}$, (e) $\mathcal{H}_{OR} = \text{OR}(\mathcal{H}_{surf-VertSE}, \mathcal{H}_{surf-HorizSE})$ (f) $\mathcal{H}_{AND} = \text{AND}(\mathcal{H}_{surf-VertSE}, \mathcal{H}_{surf-HorizSE})$

it to extract information at various levels of the hierarchies, which often requires a hard parameter-tuning [12]. Furthermore, it obliterates the interesting property that hierarchical segmentations are more informative than flat segmentations as they capture simultaneously cluster structure at all levels of granularity. The study of most of the useful existing hierarchical clustering methods is made complex by the fact that they are defined algorithmically. The morphological hierarchies that we use fall within this scope. To study the properties of such hierarchies, works define a distance between hierarchical clusterings called Gromov-Hausdorff distance [13, 14], and use it to study their stability and convergence [15]. In this work, we make use of this distance to build a feature space that capture the relative information brought by several hierarchies applied on the same image.

3.1. Gromov-Hausdorff distance

In [13] (translated in English in [14]), the authors propose a distance, called Gromov-Hausdorff (GH) distance, to measure how far two compact metric spaces are from being isometric. It gives a very useful and natural way to distinguish between metric spaces. By reducing this distance to the subclass of ultrametric spaces, we can in particular quantify the relative contributions of different hierarchical clusterings. This distance, used intensively in several fields such as phylogenetics and data mining [16], has also notably been used in image processing as a way to estimate the similarity between two points clouds [17]. Let us consider two metric spaces (X_1, u_α) and (X_2, u_β) . One supposes that we have defined two functions $f : X_1 \rightarrow X_2$ and $g : X_2 \rightarrow X_1$ that are maps from one space to the other. The GH-distance is expressed

as:

$$d_{\text{GH}}(X_1, X_2) := \frac{1}{2} \min_{f, g} \max(d_{\text{is}}(f), d_{\text{is}}(g), d_{\text{is}}(f, g)) \quad (1)$$

With the distortion $d_{\text{is}}(f)$ and the joint distortion $d_{\text{is}}(f, g)$ defined as:

$$\begin{cases} d_{\text{is}}(f) := \max_{(x, x') \in X_1^2} |u_\alpha(x, x') - u_\beta(f(x), f(x'))| \\ d_{\text{is}}(f, g) := \max_{x \in X_1, x' \in X_2} |u_\alpha(x, g(x')) - u_\beta(x', f(x))| \end{cases} \quad (2)$$

Intuitively, it measures how close can we get to an isometric (distance-preserving) embedding between two metric spaces. As it is shown in [13, 14, 15], to determine (1) for two hierarchies defined over different sets, one must match data points before any distance computation, which is a computationally heavy operation that leads some authors to provide heuristics to approximate it in specific configurations [17, 18]. In our case we are dealing with structures that are much simpler: instead of considering objects-to-objects distances, we compute them for hierarchies based upon the same fine partition of an image. So the considered metric spaces differ only by their metrics and not by the space they cover, which means that the two distortions are symmetrical and equal to the joint-distortion as well. Thus, the GH distance (1) simply becomes:

$$d_{\text{GH}}((X, u_\alpha), (X, u_\beta)) = \max_{x, x' \in X} |u_\alpha(x, x') - u_\beta(x, x')|. \quad (3)$$

3.2. Ultrametric Normalization

To make sense of GH distances, one must be sure that ultrametric values are commensurable and of the same order of magnitude. Accordingly, we propose a way to normalize these values with respect to the number of regions in each level of the hierarchy. Let (\mathcal{H}, λ) be an ultrametric hierarchy, with $\lambda : \mathcal{X} \mapsto \mathbb{R}_+$ which is strictly increasing with the inclusion order over \mathcal{H} . Let us denote $N = \text{card}(H)$, $(\mathcal{H}_0, \dots, \mathcal{H}_N)$ the nested series of clustering associated with the hierarchy, which are numbers of regions in successive levels of the hierarchy (n_0, \dots, n_N) , with $0 < n_0 \leq n_1 \leq \dots \leq n_N = N$. Then we take as a *normalized ultrametric*:

$$\tilde{\lambda} : \begin{cases} \mathcal{H} \rightarrow [0, 1] \\ \mathcal{H}_i \mapsto \frac{N - n_i}{N} \end{cases} \quad (4)$$

3.3. An interhierarchy distance matrix as hierarchical feature

Provided with such a distance, we can now quantify the relative contributions of different hierarchies built upon the same image. This provides us with a condensed representation leveraging the information provided by all the different

levels of these different hierarchies. To do so, let us consider an image \mathbf{I} and a set of complementary hierarchies $((\mathcal{H}_1, \lambda_1), \dots, (\mathcal{H}_N, \lambda_N))$ built upon this image. As we have seen in the previous section, it is then straightforward to compute the GH-distance between these hierarchies, as they constitute ultrametric spaces upon the same set. We take advantage of it by building the following symmetrical distance matrix:

$$M(\mathbf{I}, (\mathcal{H}_1, \dots, \mathcal{H}_N)) = [d_{\text{GH}}(\lambda_i, \lambda_j)]_{(i, j) \in \{1, \dots, N\}^2} \quad (5)$$

Since this matrix is symmetrical, we retain for each image its upper triangular part only. This constitutes a descriptor of the image for which we only had to specify the high-level parameters governing the hierarchies generation. Summarizing, an overview of the proposed process to extract features from a family of hierarchies can be found in Fig. 1.

4. EXPERIMENTAL RESULTS

In this section, we present some experimental results highlighting the properties of the descriptors we have proposed in the previous section.

4.1. Dead leaves process classification

In a first experiment, we highlight the discriminative power of the unsupervised hierarchical features we introduced, as long as their understandability. In the spirit of [19], we want to test if these features capture pertinent information leading to a quicker understanding of the images. To do so, we consider a classification problem on a set of simulated images from different dead leaves process [20][21]. In a dead leaves model, two dimensional textured surfaces (which are called “leaves” or “primary grains”) are sampled from a shape and size distribution and then placed on the image plane at random positions, occluding one another to produce an image. It is well-known that such a model creates images which share many properties with natural images such as scale invariance and other statistics properties [22][23]. In our experiment, we have simulated five classes with 100 images each, by using dead leaves model with different primary grains: circles, crosses, flowers, horizontal and vertical lines. Examples of simulated images can be found in Fig.3. Note that we have included different sizes and orientations tweaks to increase the difficulty of the identification. For each of these images, we compute the following hierarchies: dynamics, surface-based SWS hierarchies with structuring elements of various sizes and forms (cross, circle, diagonals, horizontal and vertical lines), as long as probabilistic AND and OR of these SWS hierarchies. Then we generate for each of these images the interhierarchy distance matrices of equation 5. We can then use these matrices as features in a classical classification pipeline using a linear support vector machines (SVM) to classify images of each class. We notice that the system can

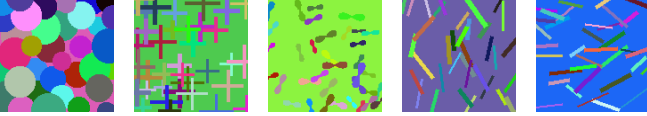


Fig. 3. False-color representation of simulated images by dead leaves model with different primary grains.

learn with very few examples how to discriminate properly these five classes. In order to have a comparison point, we conduct the same experiment using a Convolutional Neural Network (CNN) with a two-layer architecture¹ without image augmentation to make a fair comparison. In Fig. 4 are represented for both experiments the evolutions of the average f-score with the percentage of images used in the training set. In the first experiment (using the distance matrices as features), using only 5% of them (so 25 images out of 500) already leads to a 85% percentage of good classification over the remaining images, and this figure quickly goes up. In the CNN experiment, the number of required training images to get to the same results is bigger. It is thus as if, on the con-

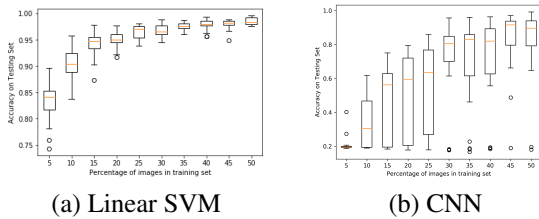
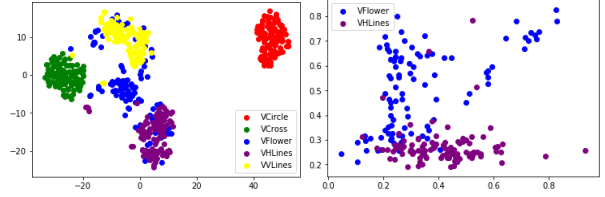


Fig. 4. Evolution of the classification error with the number of images in the training set: (a) Linear SVM on proposed features, (b) CNN.

trary to CNN that have a black-box behavior, our approach shows what is often referred to as an “aha moment”, i.e. a moment of sudden realization and comprehension [19]. This translates a form of understanding of the content of the image, which is corroborated by the study of the importance of which specific interhierarchy distances were the more useful to discriminate between two types of classes. For example, discriminating between horizontal and vertical lines will mainly be due to $d_{GH}(\mathcal{H}_{surf-VertSE}, \mathcal{H}_{surf-HorizSE})$, while discriminating between crosses and circles will mainly be due to $d_{GH}(\mathcal{H}_{surf-CrossSE}, \mathcal{H}_{surf-HexSE})$. A visualization of the quality of the feature space thus generated can be found in Fig. 5(a), where we project the features in a space of two dimensions using the t -SNE algorithm [24]. Furthermore, using the variable selection method L_1 -SVM

¹(12 Conv + 12 Conv + MaxPolling(3×3) + Dropout(0.3)) + (24 filters + 24 filters + MaxPolling(3×3) + Dropout(.5)) + FullyCon64 + Dropout(.5) + SoftMax. Categorical cross-entropy as loss function and adaptive gradient (Adagrad) as optimizer. The code used to train the CNN and scripts to generate the synthetic images are available online at [https://github.com/.](https://github.com/)



(a)2D scatterplot by t -SNE(b)SD scatterplot by L_1 -SVM

Fig. 5. (a) We notice that the classes “Flowers” and “Horizontal Lines” are not well separated (b) These two distances between hierarchies provide a geometrical understanding of the images content.

[25], we can isolate the more discriminative distances for two specific classes to separate. For example, the t -SNE visualization in Fig. 5(a) shows us that discriminating between the classes “Flowers” and “Horizontal Lines” is not straightforward. The more discriminative variable between these two classes is the distance between $\mathcal{H}_{surf-VertSE}$ and $\mathcal{H}_{AND(surf-VertSE, surf-HexSE)}$: this is a geometrical interpretation of the image content, as they respectively capture straight lines and lines with a protuberance (i.e. flowers). Projecting the distance features onto the subspace of the two more discriminative variables properly separates these two classes, as can be seen in Fig. 5(b). We are thus provided with a good way to control and understand features generation to be further used in vision-based system, especially when we have prior information about the type of objects we are looking for in images, and/or when we have very few training examples and want to maximize their usefulness.

5. CONCLUSIONS AND DISCUSSIONS

In this work, we have proven the efficiency of features created by estimating the differences between different hierarchical segmentations of the same images. The proof-of-concept experiments we conducted show that these features can provide a geometrical interpretation of the image content, and help to analyze images classes with very few examples. Furthermore, this approach can be extended to any type of hierarchies to capture various types of information. We see three ways to use such information. First, if we know well the type of images we want to discriminate, we can generate at will the hierarchies that will be the more discriminative. Secondly, if we build a sufficiently wide range of complementary hierarchies for an unknown class of images, we can characterize these images by looking at the way they react to these operators through the study of interhierarchy distances. In line with classical morphological approaches, we can thus gain knowledge about images by studying the way they react to given operators. Finally, many methods exist to extract a segmentation out of a hierarchical segmentation, and our approach provides a way to identify the hierarchies that will lead to the best results in such a process.

6. REFERENCES

- [1] S. Beucher, “Watershed, hierarchical segmentation and waterfall algorithm,” in *Mathematical morphology and its applications to image processing*, pp. 69–76. Springer, 1994.
- [2] F. Meyer, “The dynamics of minima and contours,” in *Mathematical Morphology and its Applications to Image and Signal Processing*, pp. 329–336. Springer, 1996.
- [3] L. Najman and M. Schmitt, “Geodesic saliency of watershed contours and hierarchical segmentation,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 18, no. 12, pp. 1163–1173, 1996.
- [4] G. K. Ouzounis and M. H.F. Wilkinson, “Mask-based second-generation connectivity and attribute filters,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 29, no. 6, pp. 990–1004, 2007.
- [5] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 33, no. 5, pp. 898–916, 2011.
- [6] F. Meyer, “Stochastic watershed hierarchies,” in *Eighth International Conference on Advances in Pattern Recognition*. IEEE, 2015, pp. 1–8.
- [7] J. Cousty, L. Najman, Y. Kenmochi, and S. Guimaraes, “Hierarchical segmentations with graphs: quasi-flat zones, minimum spanning trees, and saliency maps,” *Journal of Mathematical Imaging and Vision*, pp. 1–24, Oct. 2017.
- [8] A. Fehri, S. Velasco-Forero, and F. Meyer, “Automatic selection of stochastic watershed hierarchies,” in *24th European Signal Processing Conference*. IEEE, 2016, pp. 1877–1881.
- [9] J. Angulo and D. Jeulin, “Stochastic watershed segmentation,” in *8th International Symp. on Mathematical Morphology*, 2007, pp. 265–276.
- [10] A. Fehri, S. Velasco-Forero, and F. Meyer, “Prior-based hierarchical segmentation highlighting structures of interest,” in *International Symp. on Math. Morphology and Its Applications to Signal and Image Processing*. Springer, 2017, pp. 146–158.
- [11] D. S. Maia, A. de Albuquerque Araujo, Cousty J., Najman L., Perret B., and Talbot H., “Evaluation of combinations of watershed hierarchies,” in *International Symp. on Math. Morphology and Its Applications to Signal and Image Processing*. Springer, 2017, pp. 133–145.
- [12] L. Guigues, J. P. Cocquerez, and H. Le Men, “Scale-sets image analysis,” *International Journal of Computer Vision*, vol. 68, no. 3, pp. 289–317, 2006.
- [13] M. Gromov, J. Lafontaine, and P. Pansu, *Structures métriques pour les variétés riemanniennes*, Cedic, 1981.
- [14] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Springer Science & Business Media, 2007.
- [15] F. Carlsson, G. and Mémoli, “Characterization, stability and convergence of hierarchical clustering methods,” *Journal of machine learning research*, vol. 11, no. Apr, pp. 1425–1470, 2010.
- [16] J. Felsenstein, *Inferring phylogenies*, vol. 2, Sinauer associates Sunderland, MA, 2014.
- [17] G. Mémoli, F. and Sapiro, “Comparing point clouds,” in *ACM SIGGRAPH symposium on Geometry processing*. ACM, 2004, pp. 32–40.
- [18] Pankaj K Agarwal, Kyle Fox, Abhinandan Nath, Anastasios Sidiropoulos, and Yusu Wang, “Computing the gromov-hausdorff distance for metric trees,” in *International Symposium on Algorithms and Computation*. Springer, 2015, pp. 529–540.
- [19] Z. Yan and X. S. Zhou, “How intelligent are convolutional neural networks?,” *arXiv preprint arXiv:1709.06126*, 2017.
- [20] D. Jeulin, “Dead leaves models: From space tessellation to random functions,” in *Advances in Theory and Applications of Random Sets*, pp. 137–156. World Scientific Publishing, 1997.
- [21] Georges Matheron, *Random sets and integral geometry*, Wiley New York, 1975.
- [22] D. Zoran and Y. Weiss, “Natural images, gaussian mixtures and dead leaves,” in *Advances in NIPS*, 2012, pp. 1736–1744.
- [23] X. Pitkow, “Exact feature probabilities in images with occlusion,” *Journal of vision*, vol. 10, no. 14, pp. 42–42, 2010.
- [24] Laurens van der Maaten and Geoffrey Hinton, “Visualizing data using t-SNE,” *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.
- [25] Ji Zhu, Saharon Rosset, Robert Tibshirani, and Trevor J Hastie, “1-norm support vector machines,” in *Advances in NIPS*, 2004, pp. 49–56.