Texture Based Classification Of Seismic Image Patches Using Topological Data Analysis

Rahul Sarkar* & Bradley Nelson*†

*Institute for Computational and Mathematical Engineering, Stanford University, †Stanford Linear Accelerator

Abstract

We have developed a supervised learning method for textural classification of seismic image patches, based on a topological tool called *persistent homology.* For each image, persistent homology produces a list of birth-death pairs which describe how the topology of the image changes as a function of pixel values. Feature vectors are extracted from these pairs, which are in turn used to train machine learning classifiers for the problem. In addition, we study the efficacy of different derived textural attributes when used in place of the raw images in the workflow. Our proposed method is tested on the publicly available LANDMASS datasets, and our results indicate that these features can be quite effective at capturing qualitative textural information in seismic images.

Introduction

Seismic image segmentation is an important activity in hydrocarbon exploration, which to a large extent is a manual task performed by human interpreters. Development of seismic attributes to aid interpreters was an important milestone, and more recently their combination with machine learning (ML) algorithms has led to partial automation of this process. A subclass of these methods attempts to perform image segmentation directly using textural attributes [2]. For example, a salt body looks very different from a region of sedimentary deposits in a seismic image. An important recent work by [3] demonstrates that seismic image segmentation can be performed by breaking up the image into small patches, followed by using convolutional neural networks (CNNs) to classify these patches based on their textures.

Our method uses a similar patch-based approach to seismic image classification, using features derived from topological data analysis (TDA)[4]. These features are theoretically robust to challenges in texture recognition such as image rotation, scaling, nonlinear deformation, and pixel value perturbations. In addition to the raw images, we produce images derived from the raw images using gray-level co-occurrence matrices (GLCM)[5], and experiment with them as inputs to our workflow.

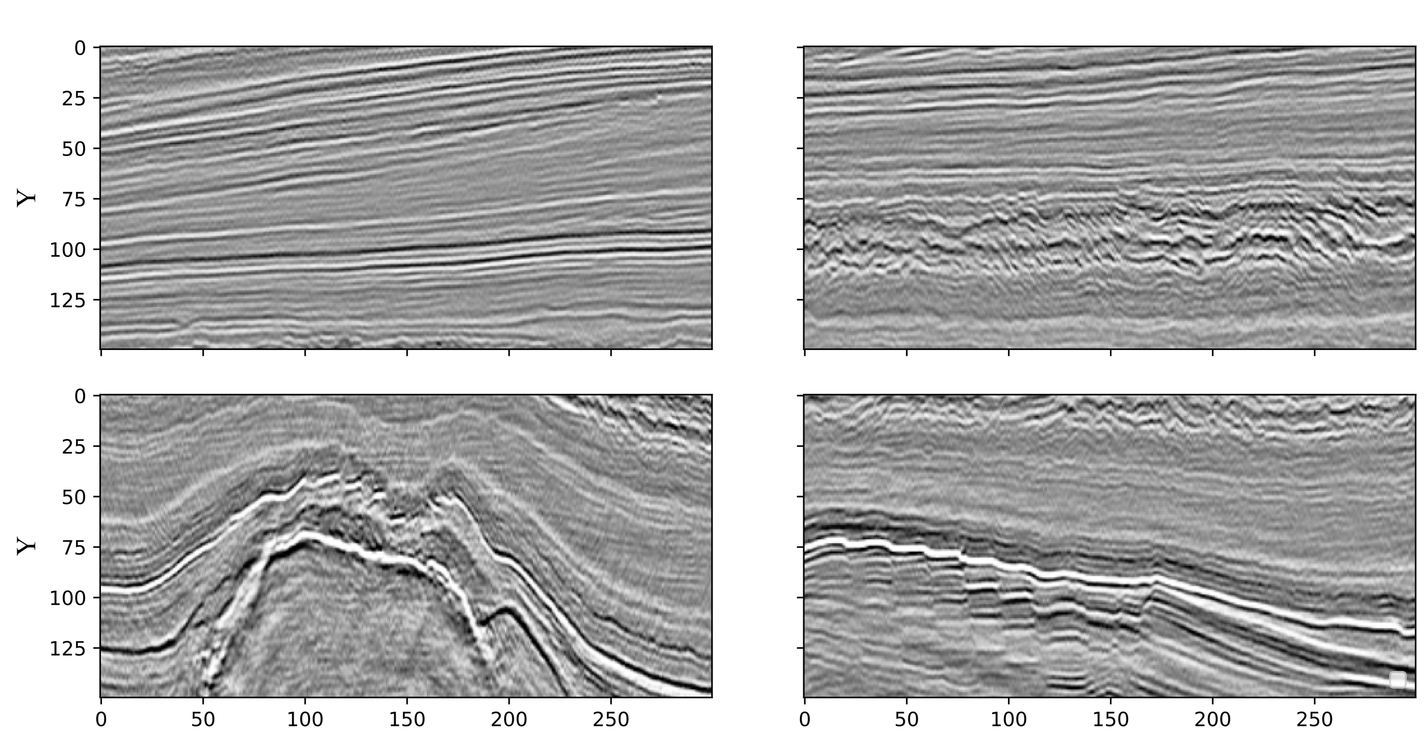

Figure: Example images from the LANDMASS-2 data set. Clockwise from top left: flat horizon, chaotic horizon, fault, salt body.

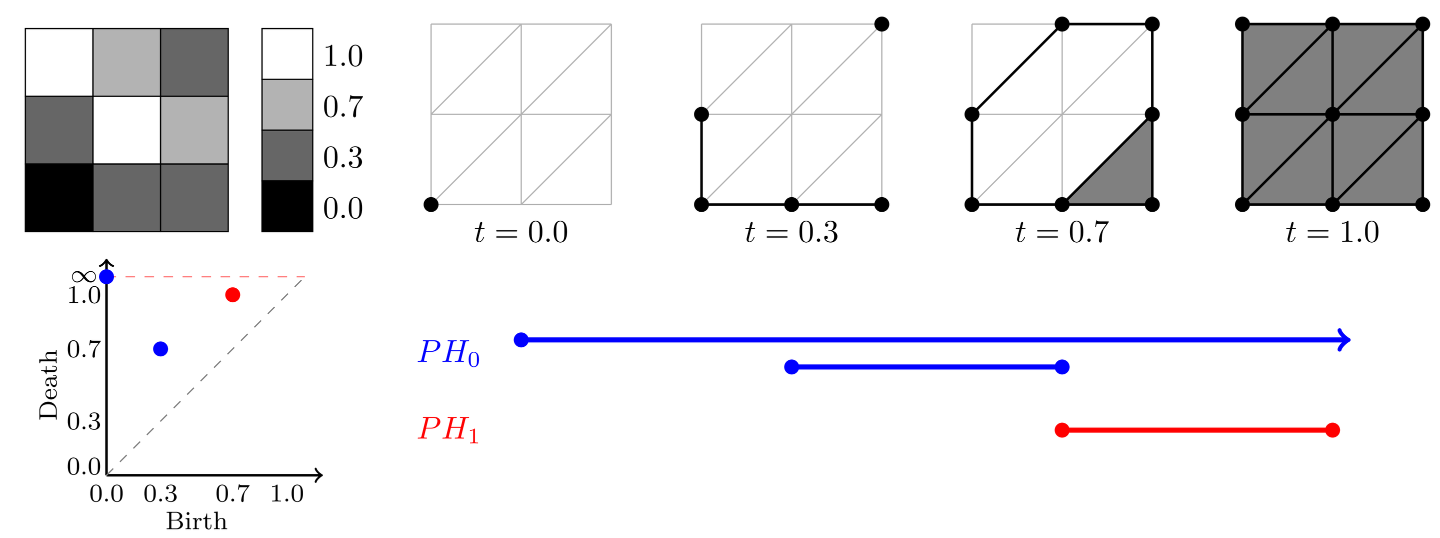

Figure: Example sub-level set filtration on a 3×3 image. Top left: image. Top right: filtration at homological critical points. Bottom right: persistence barcode. Bottom left: persistence diagram.

Topological Features

Images may become topological spaces by considering pixels as points, and using the Freudenthal triangulation of the lattice to add edges to adjacent points and triangles to fill in 2×2 pixel patches. Interesting information appears by creating a filtration on the space by restricting the set of active vertices to pixels that have value $\leq t$. Persistent homology tracks how connected components and holes appear and disappear in the filtration as the filtration parameter increases. The result is a collection of birth-death pairs: PH_0 (for connected components) and PH_1 (for holes). These collections of points can be visualized using persistence diagrams, which plot the pairs in the plane.

Each image may produce a different number of birth-death pairs. In order to produce a standard set of features, we use a set of polynomial functions [6]

$$p(\alpha; \{(b_i, d_i)\}_{i \in J}) = \frac{1}{|J|} \sum_{i \in J} \sum_{j,k} \alpha_{j,k} (d_i - b_i)^j (d_i + b_i)^k$$

We use $\alpha = \delta_{j,k}, (j,k) \in \{0,1,2,3\}^2 - (0,0)$ for homology dimensions 0 and 1. This produces 30 features in total. This can be thought of as taking integrals of functions over the point measures supported on the birth-death pairs.

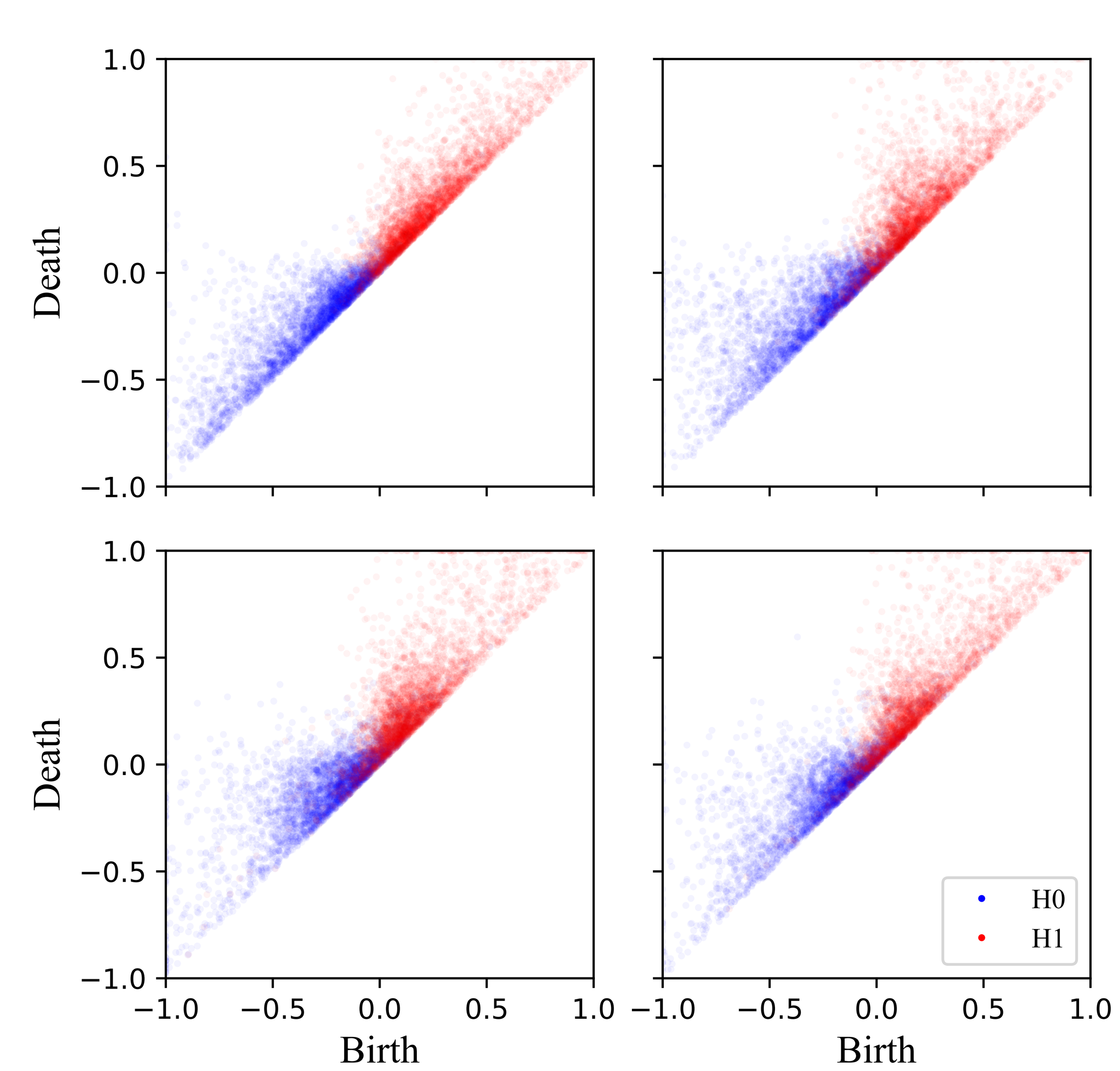

Figure: Persistence diagrams produced from the example images in the LANDMASS-2 data set.

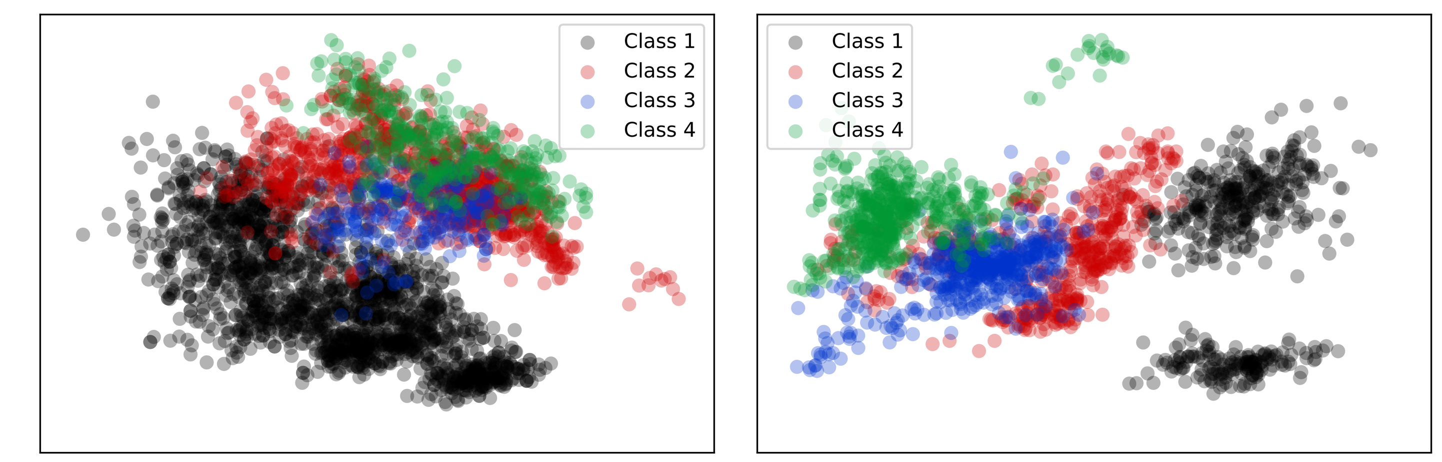

Figure: Left: PCA embedding of LANDMASS-1 data set using topological features. Right: PCA embedding of LANDMASS-2 data set using topological features

Results

We test our method on the publicly available LANDMASS data sets [7]. We use the 30 topological features from each image, and train three different black-box classification algorithms: a multi-class support vector machine (SVM), a random forest (RF), and a simple neural network (NN). We also use GLCM features to produce additional persistence diagrams.

Attribute	Classification Accuracy on Test Set (%)		
	SVM	RF	NN
Raw Image	99.8 / 75.2 / 0.0 / 0.0 100.0 / 55.0 / 88.3 / 74.3	99.9 / 98.6 / 95.2 / 93.3 100.0 / 98.0 / 100.0 / 96.3	100.0 / 99.6 / 99.7 / 98.4 100.0 / 100.0 / 99.0 / 95.0
GLCM Mean	100.0 / 18.6 / 34.1 / 29.3 62.7 / 19.0 / 4.0 / 100.0	99.9 / 97.9 / 82.1 / 93.3 100.0 / 97.0 / 97.3 / 91.7	100.0 / 97.8 / 92.8 / 97.0 100.0 / 96.0 / 95.7 / 96.3
RMS Amplitude	100.0 / 1.0 / 0.0 / 0.0 74.7 / 85.7 / 71.3 / 61.7	99.3 / 96.1 / 88.0 / 82.0 99.7 / 96.0 / 96.0 / 91.7	99.5 / 99.1 / 96.3 / 91.5 99.7 / 99.0 / 93.7 / 91.3
GLCM Correlation	100.0 / 0.0 / 0.0 / 0.0 64.7 / 32.0 / 89.3 / 32.3	99.3 / 94.9 / 80.8 / 91.2 99.7 / 93.7 / 92.0 / 97.0	99.8 / 93.6 / 87.7 / 96.7 100.0 / 95.7 / 93.7 / 98.3
GLCM Variance	96.6 / 94.1 / 92.8 / 67.7 97.3 / 93.3 / 91.7 / 87.0	98.5 / 95.7 / 96.3 / 74.0 99.0 / 95.3 / 96.7 / 89.7	99.3 / 98.3 / 98.1 / 87.3 99.7 / 99.0 / 99.3 / 95.0

References

[1] Rahul Sarkar and Bradley Nelson. Texture based classification of seismic image patches using topological data analysis. In *81st EAGE Conference and Exhibition*, 2019. forthcoming.

[2] PL Love and M Simaan. Segmentation of stacked seismic data by the classification of image texture. In *SEG Technical Program Expanded Abstracts 1984*, pages 480–482. Society of Exploration Geophysicists, 1984.

[3] Daniel Salles Chevitarese, Daniela Szwarcman, Emilio Vital Brazil, and Bianca Zadrozny. Efficient classification of seismic textures. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[4] Gunnar Carlsson. Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308, 2009.

[5] Christoph Georg Eichkitz, Johannes Amtmann, and Marcellus Gregor Schreilechner. Calculation of grey level co-occurrence matrix-based seismic attributes in three dimensions. *Computers & Geosciences*, 60:176–183, 2013.

[6] Aaron Adcock, Erik Carlsson, and Gunnar Carlsson. The ring of algebraic functions on persistence bar codes. *Homology, Homotopy and Applications*, 18(1):381–402, 2016.

[7] Yazeed Alaudah, Zhen Wang, Zhiling Long, and Ghassan AlRegib. Landmass seismic dataset. `http://cegp.ece.gatech.edu/codedata/LANDMASS/index.html`, 2015.

Acknowledgments

We would like to thank Biondo Biondi and Gunnar Carlsson for providing us guidance on this project, and for many helpful suggestions. R.S. was partially supported by the Stanford Exploration Project, and B.N. was partially supported by the US DoD NDSEG fellowship program, and partially supported by DOE Contract DE-AC02-76SF00515