Topological Data Analysis for Parkinson's Disease Classification and Severity Assessment

Afra Nawar, Farhan Rahman, Narayanan Krishnamurthi, Anirudh Som, Pavan Turaga

EMBC 2020 Montreal

Ira A. Fulton Schools of Engineering
Arizona State University

Geometric Media Lab

Parkinson’s Disease

Statistics

- 2nd most common neurodegenerative disease
- Incidence increasing
- 1 mil. affected in U.S., 10 mil. worldwide
- Elderly most commonly affected

Symptoms

- Muscle rigidity
- Lack of balance and postural control
- Tremors
- Lack of fine motor skills

Parkinson's Disease

UPDRS

- Unified Parkinson's Disease Rating Scale
- Includes sections based on clinical motor evaluation, self evaluation, and Mentation/Behavior/Mood

Proposed Method

- Signal Processing paired with Machine Learning (Automated Assessment Tool)

Increased Accuracy & Early Detection

Increased Objectivity

Dataset

Data Collection

Screen

Pressure
Plate
Platform

Figure 1: Patient Standing on Pressure Plate Platform

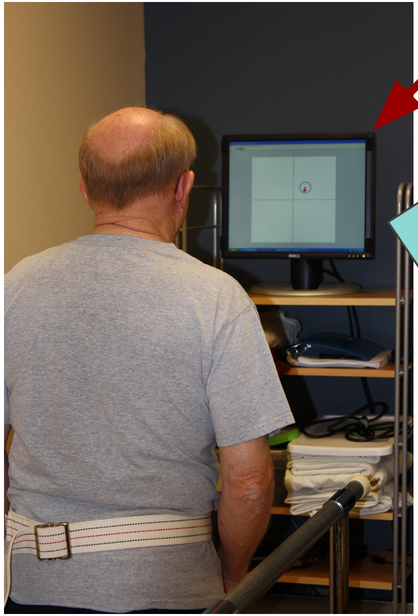


Figure 2: Patient View of screen marking targets and COP

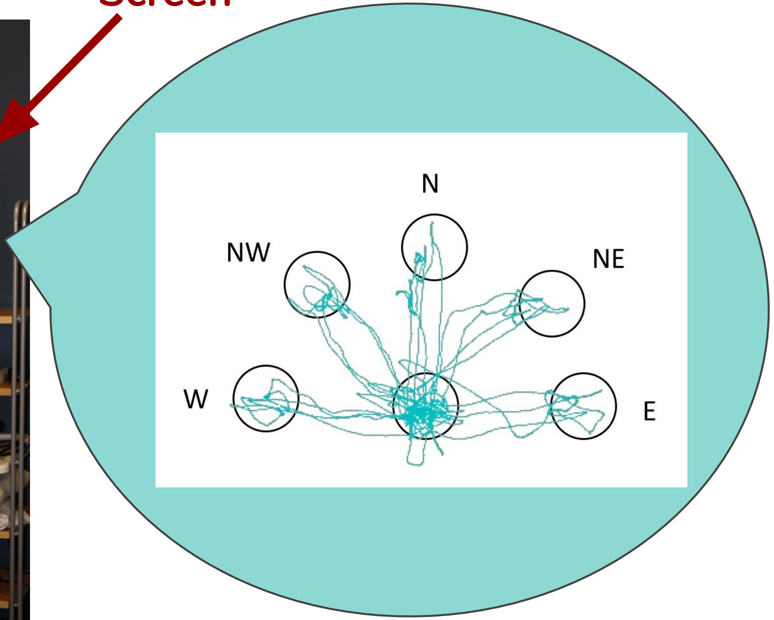


Figure 3: Visualization of COP movement during target reaches

[1] A. Som, N. Krishnamurthi, V. Venkataraman, and P. Turaga. Attractor-shape descriptors for balance impairment assessment in parkinson's disease. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 3096–3100. IEEE, 2016.
[2] N. Krishnamurthi, S. Mulligan, P. Mahant, J. Samanta, and J. J. Abbas. Deep brain stimulation amplitude alters posture shift velocity in parkinson's disease. Cognitive neurodynamics, 6(4):325–332, 2012.

Dataset Composition

Healthy Subjects

Elderly Subjects
- 22 Subjects
- 5 Trials Each

Young Subjects
- 21 subjects
- 5 Trials Each

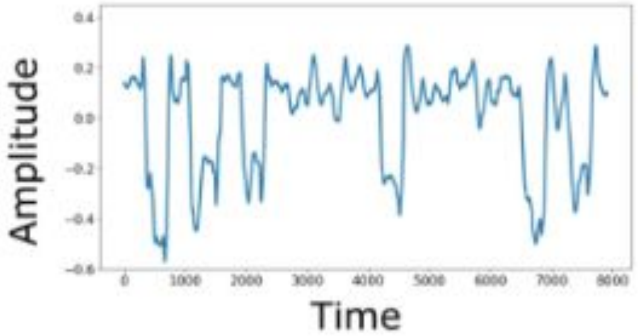
Parkinson's Subjects

Parkinson's Disease Subjects
- 17 subjects
- 3 Trials Each

60 Subjects 266 Trials

Data Representation

- X : measurement in the medio-lateral direction
- Y: measurement in the antero-posterior direction
- F_x, F_y, F_z : Forces and in the X, Y, Z direction
- M_x, M_y, M_z : Moments in the X, Y, Z direction



Each signal consists of the value in the chosen direction with respect to time.

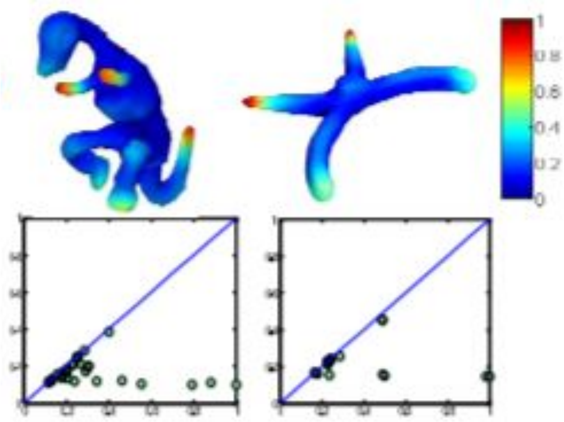
Methods

Topological Data Analysis

- Analyze Structure and Shape of Data
- Time Series, Point Clouds, Images

Why Topological Data Analysis?

- Invariant to order of target reaches
- Invariant to smooth deformations (stretching, bending, scaling, etc)

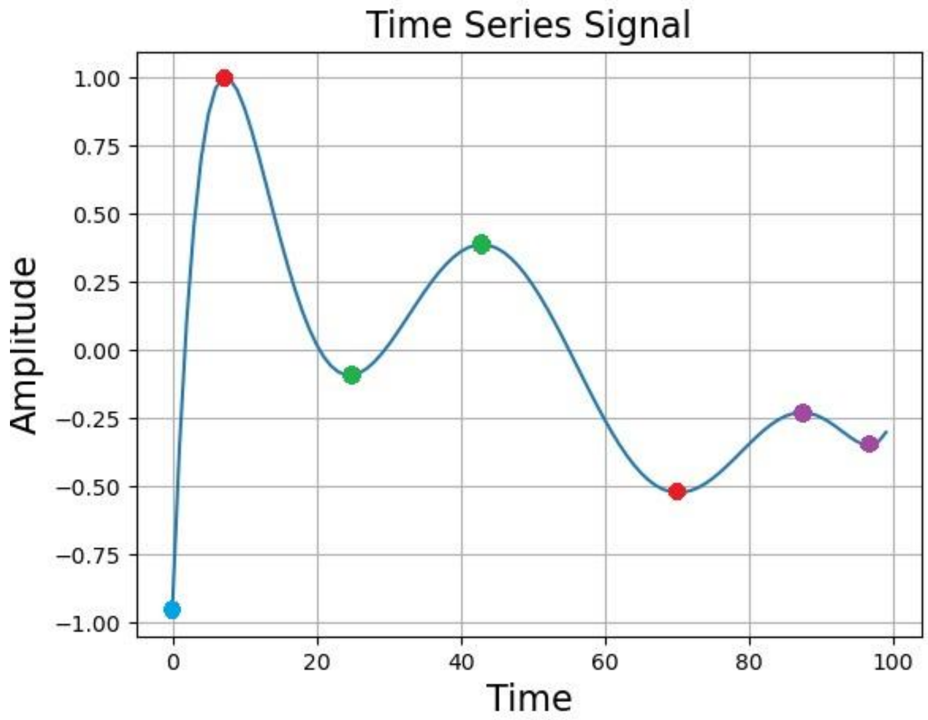


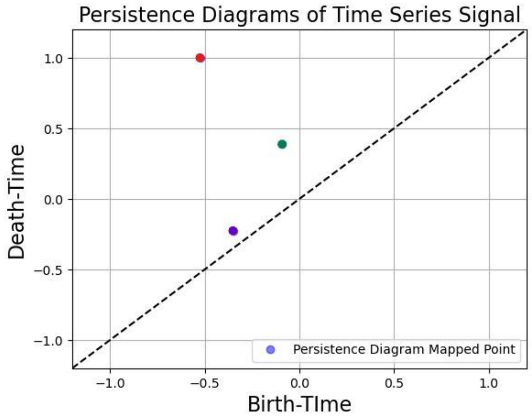
[3] C. Epstein, G. Carlsson, and H. Edelsbrunner. Topological data analysis. Inverse Problems, 27(12):120201, 2011

[4] G. Carlsson. Topology and data. Bulletin of the American Mathematical Society, 46(2):255–308, 2009.

[5] C. Li, M. Ovsjanikov, and F. Chazal. Persistence-based structural recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1995–2002, 2014

Topological Data Analysis

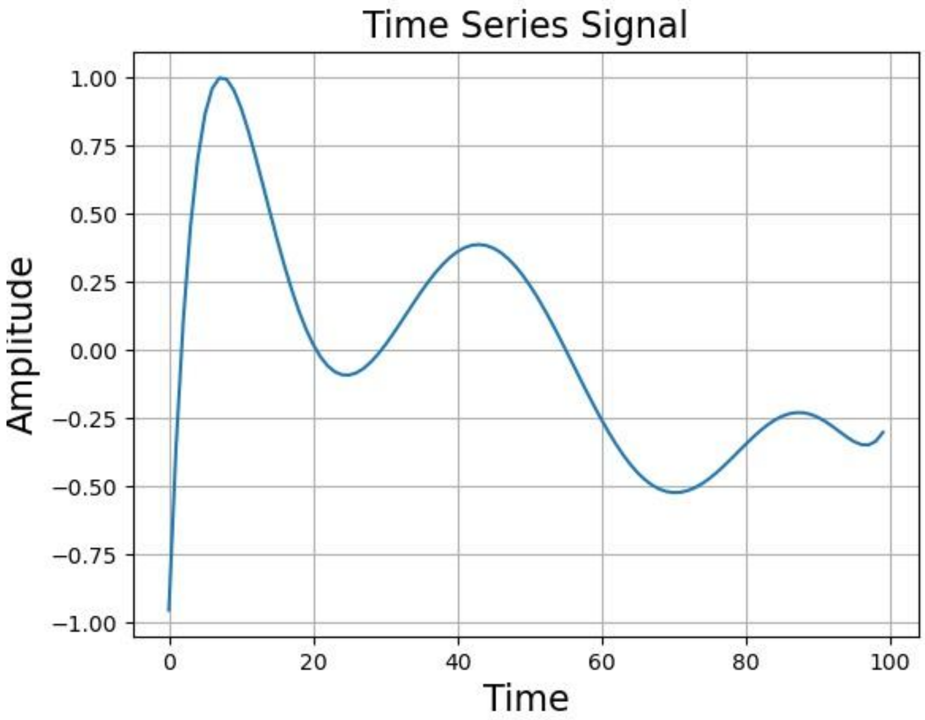


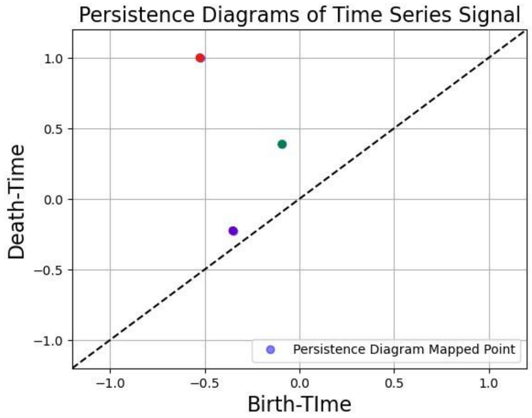


Mapping of time-series data as a pairing of its peaks with its troughs.

Represents information about the smoothness of the data, as well as its rises and falls

Topological Data Analysis

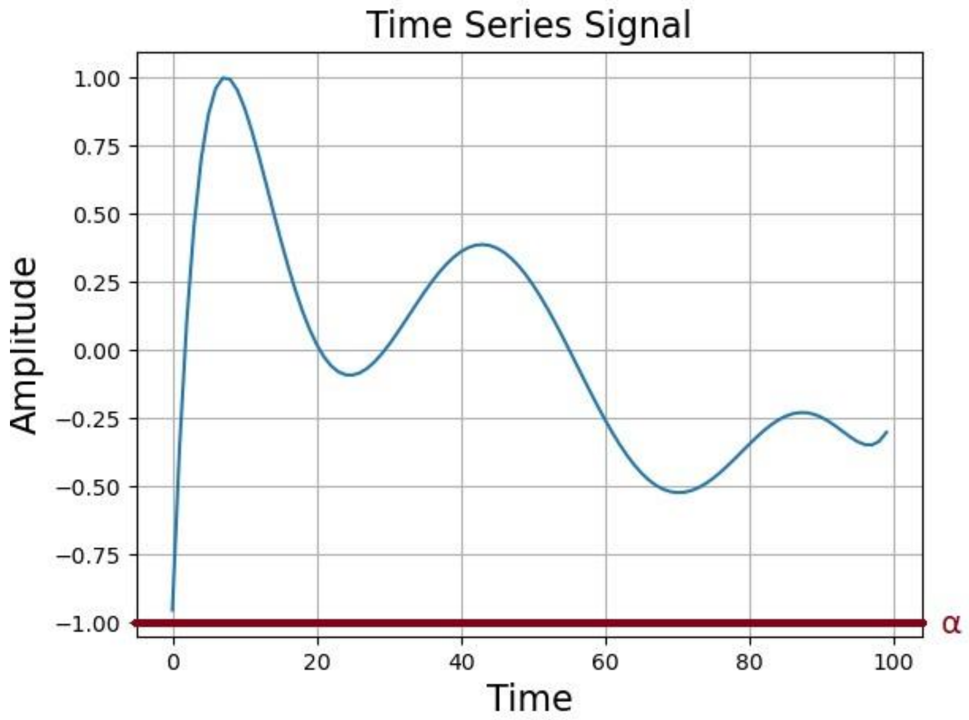


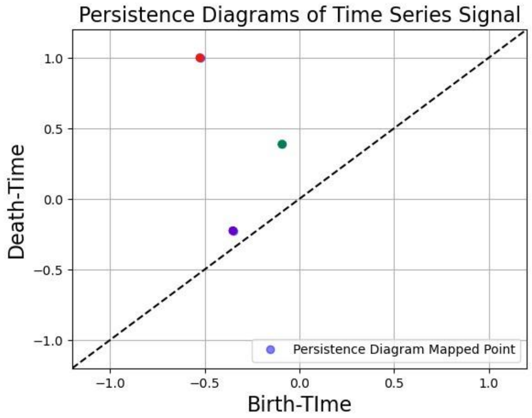


Time-series signal can be modeled as a graph with multiple nodes, each connected to two neighbors (except the ends)

Each node's value is its y-value on the plot

Topological Data Analysis

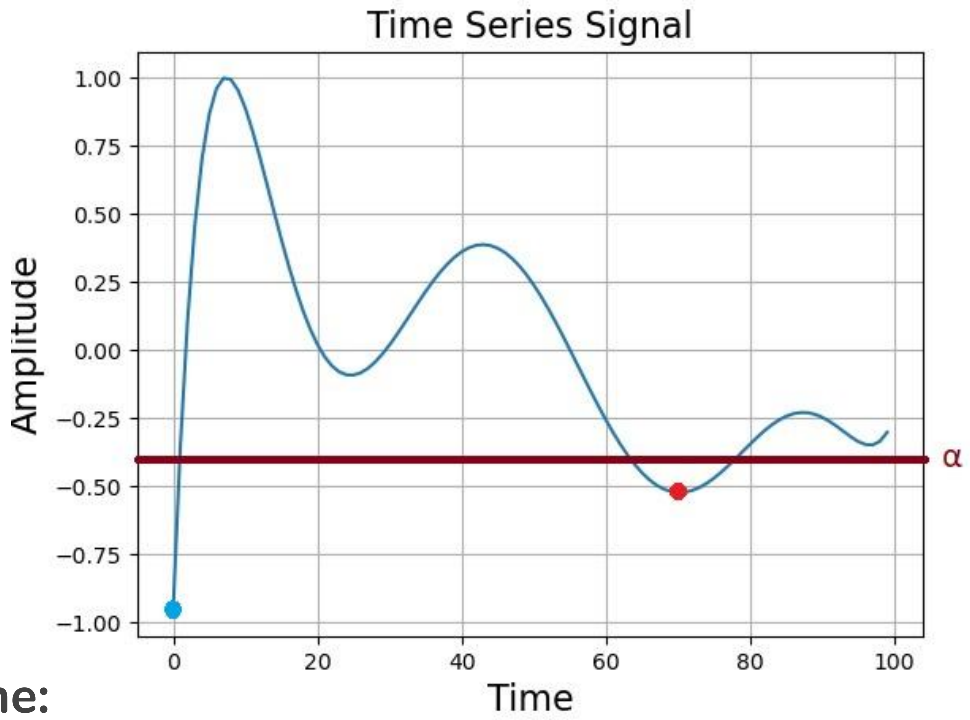


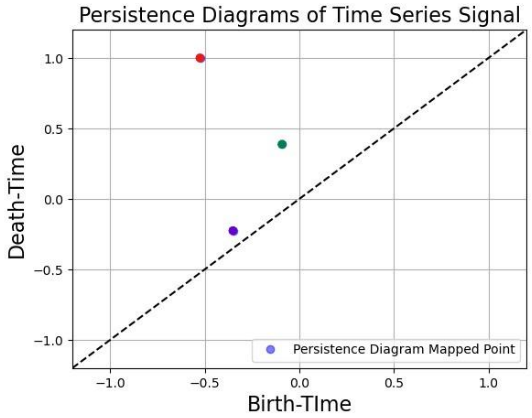


Alpha value is swept from -∞ to ∞, identifying troughs and matching them to peaks as it increases.

Local minimum (trough) is any node whose value is less than those of all of its neighbors.

Topological Data Analysis





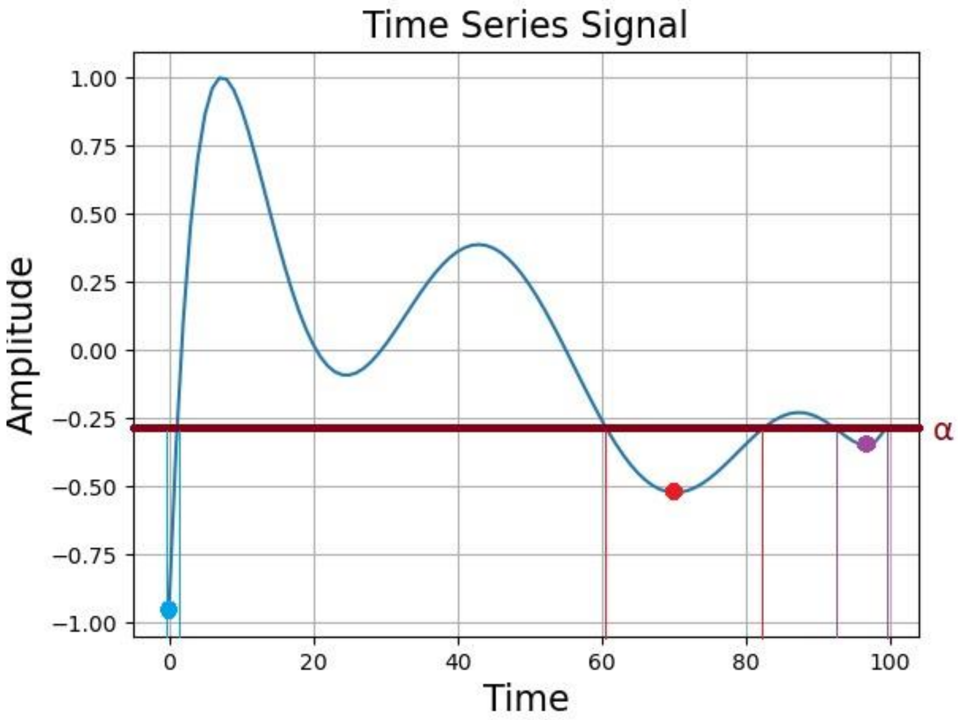
When alpha is swept passed a node that it identifies as a local minimum, it saves the value of that node as the birth of that trough

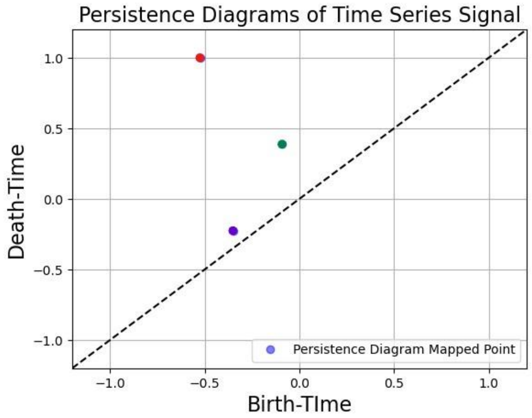
Blue trough is born at about -0.95

Red trough is born at about -0.5

Birth-time:
Value (Y-axis) of local minimum.

Topological Data Analysis



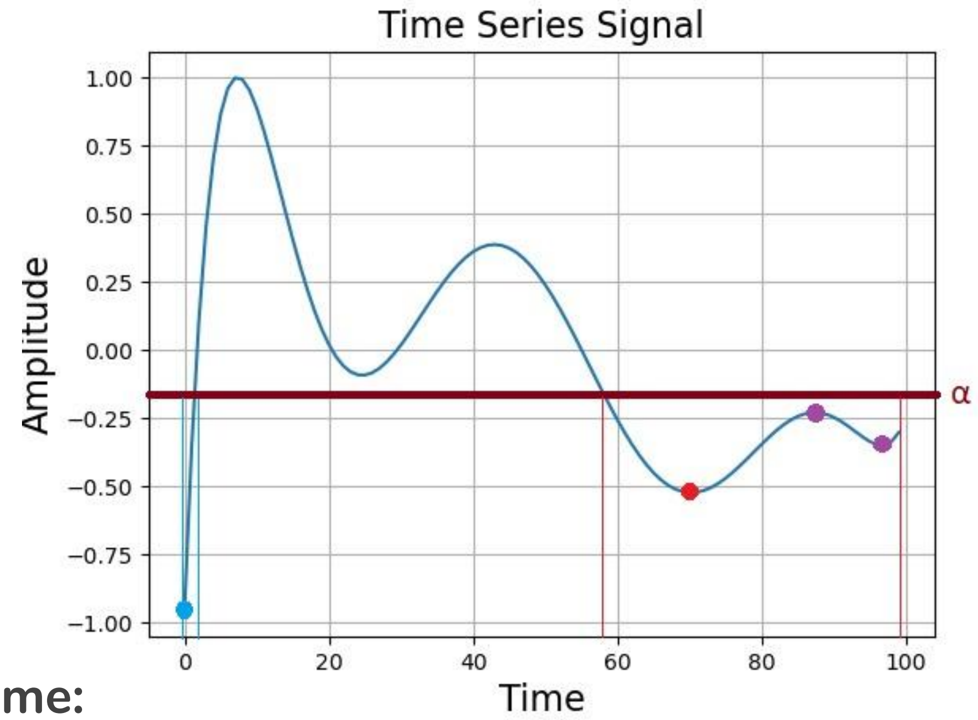


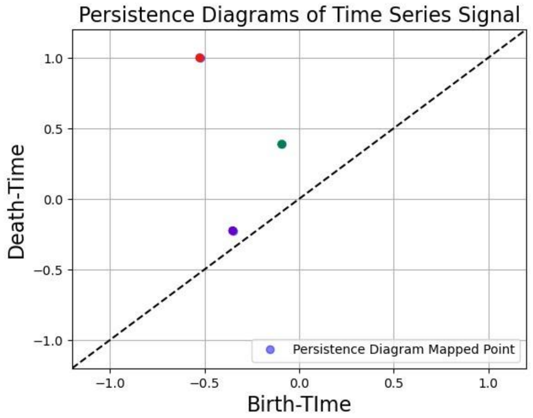
Purple trough born at about -0.3

Algorithm keeps track of the subgraphs of the overall graph for values below alpha

Alpha value splits graph into several explored subgraphs, boundaries show the minimum for each subgraph

Topological Data Analysis





The death of a trough is given by the lowest alpha value such that the trough is no longer the minimum of its subgraph.

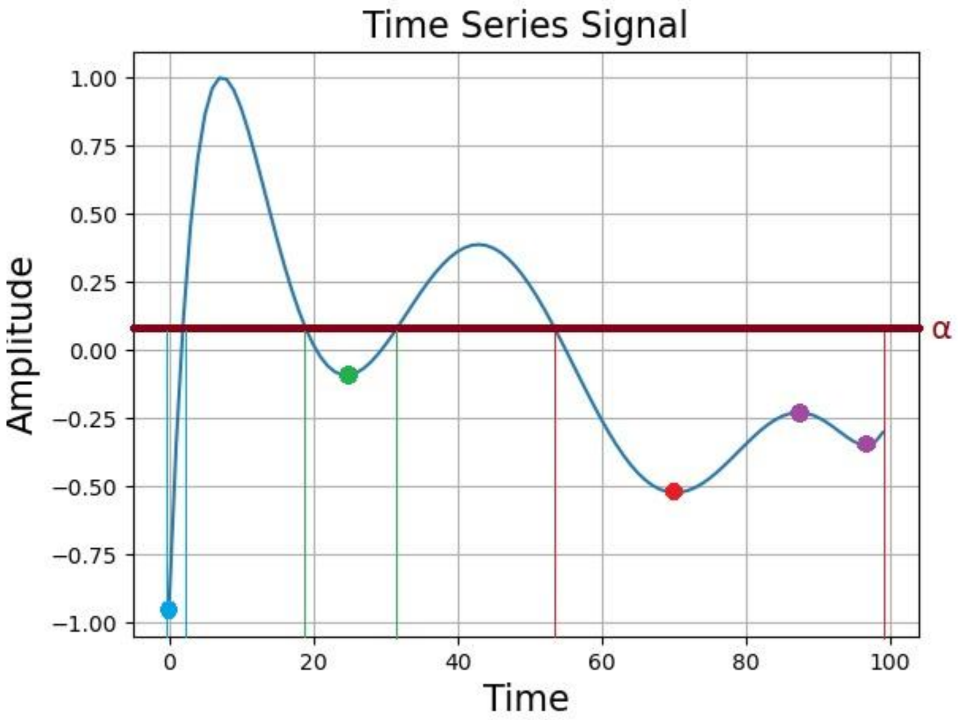
Red trough is now the minimum of the right subgraph.

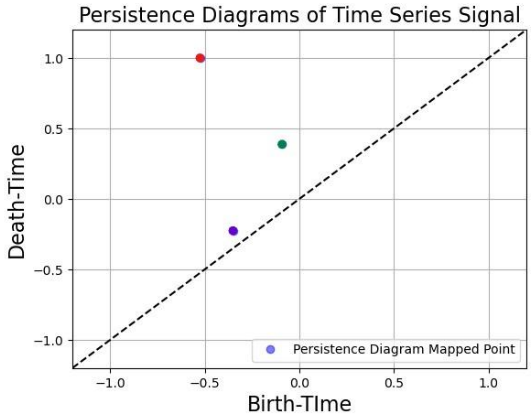
Purple trough dies at about -0.25

Death-time:

Value (Y-axis) of matched peak.

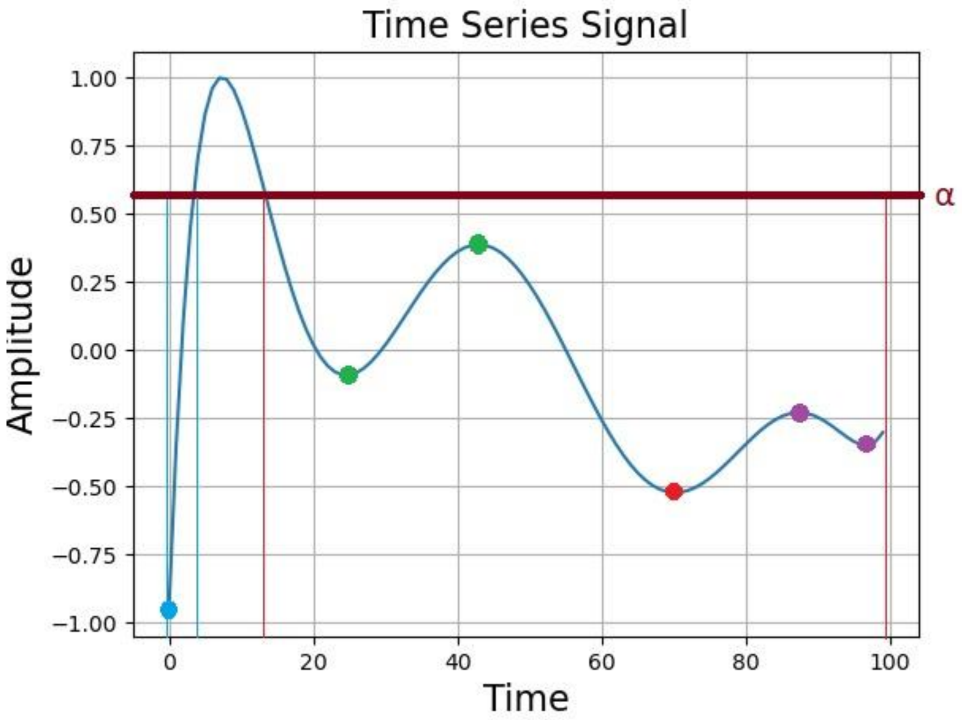
Topological Data Analysis

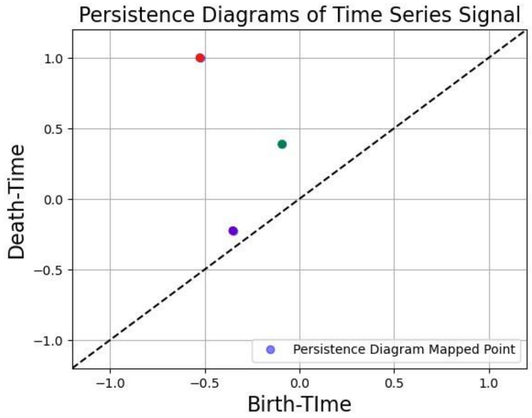




Green trough is born at about -0.1

Topological Data Analysis

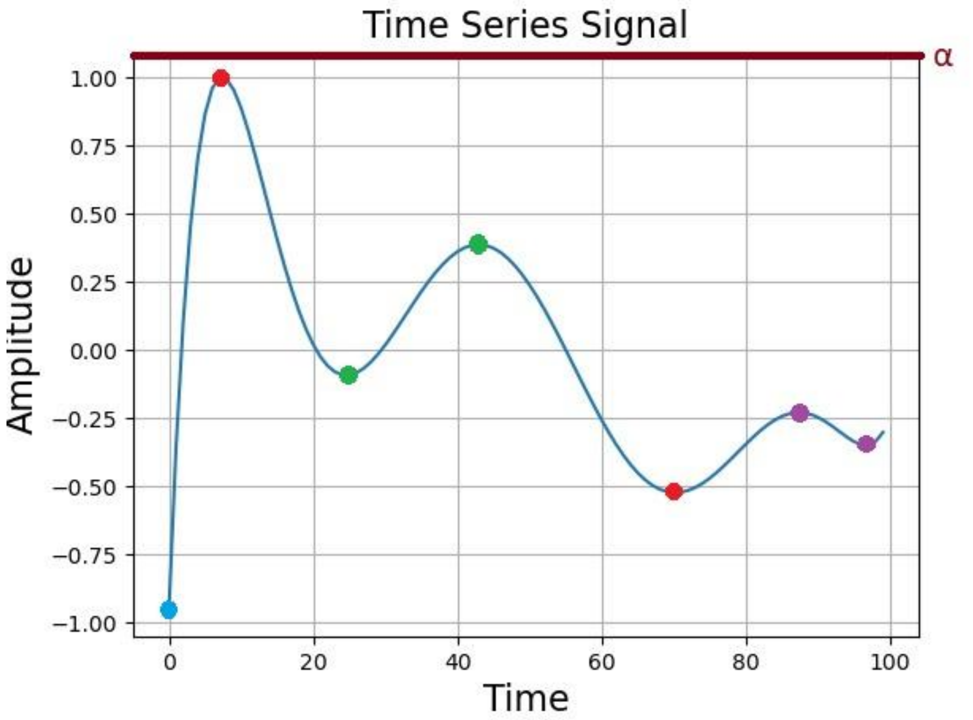


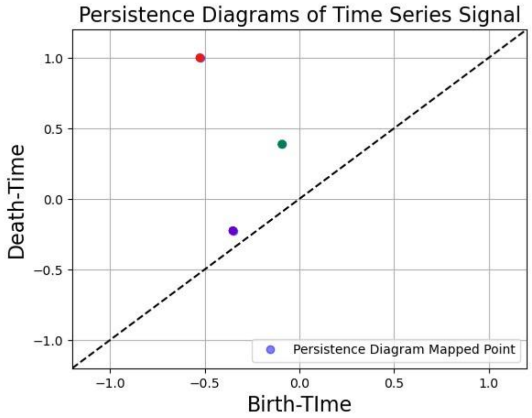


Green trough dies at about 0.35

Red trough is now the minimum of right-most subgraph.

Topological Data Analysis



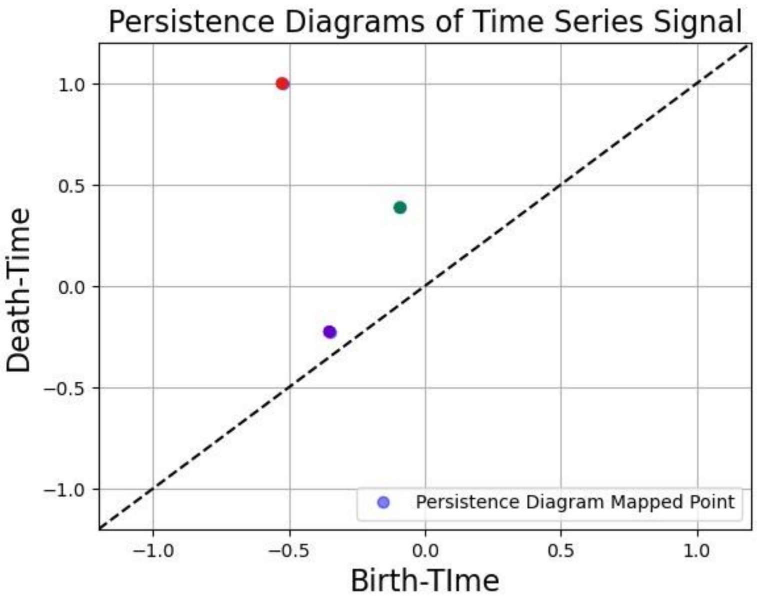


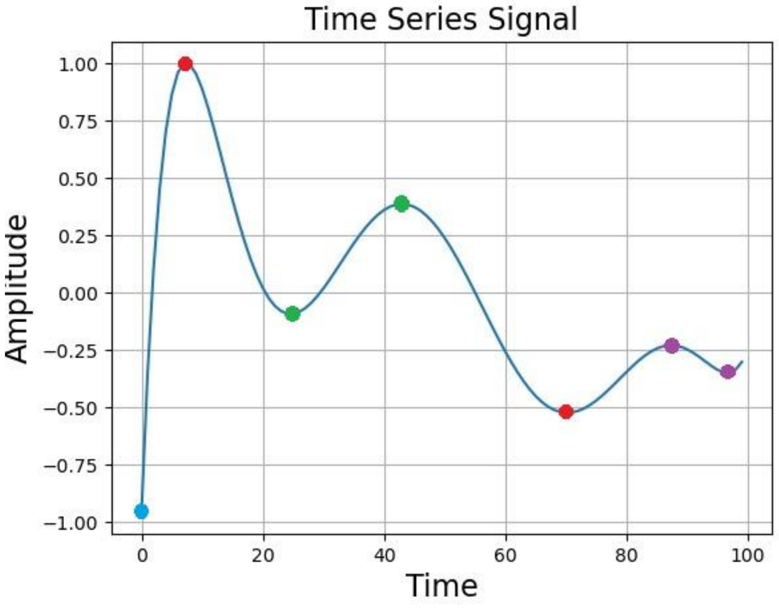
Red trough dies at about 1.

The algorithm terminates when all node values are smaller than alpha.

As the global minimum (blue trough) does not die, it is given a death time of infinity.

Persistence Diagrams





Lifetime = Death-time - Birth-time

[4] G. Carlsson. Topology and data. Bulletin of the American Mathematical Society, 46(2):255–308, 2009.

Persistence Diagrams

Pros:

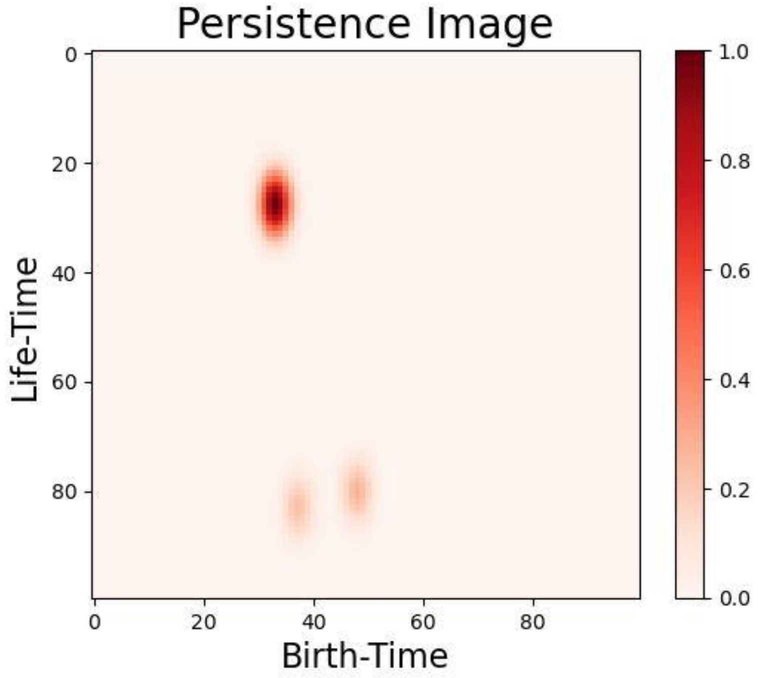
Good way to uniformly represent variable length time-series data.

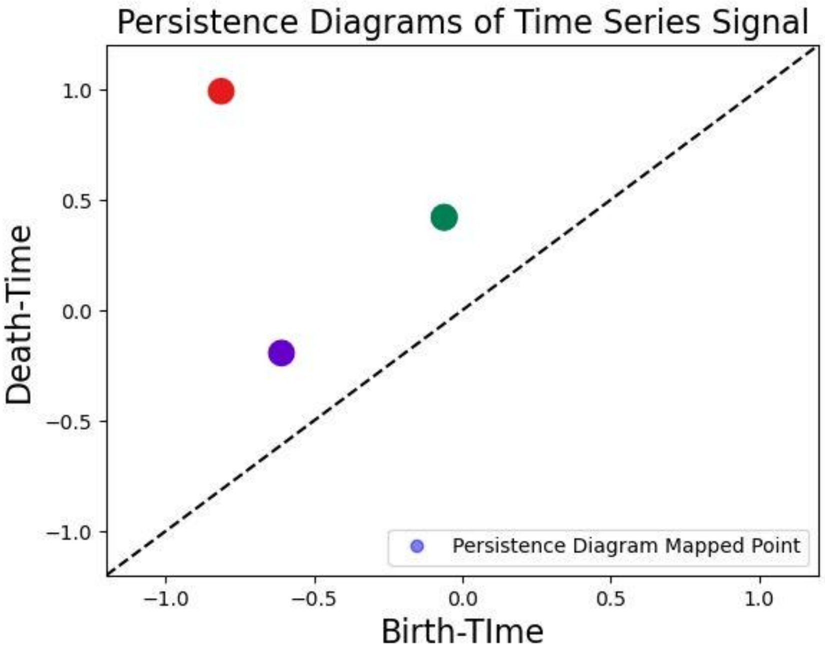
One of the most commonly used topological descriptors

Cons:

Not easily used in standard machine learning algorithms, as the number of points that we observe on a persistence diagram varies from sample to sample

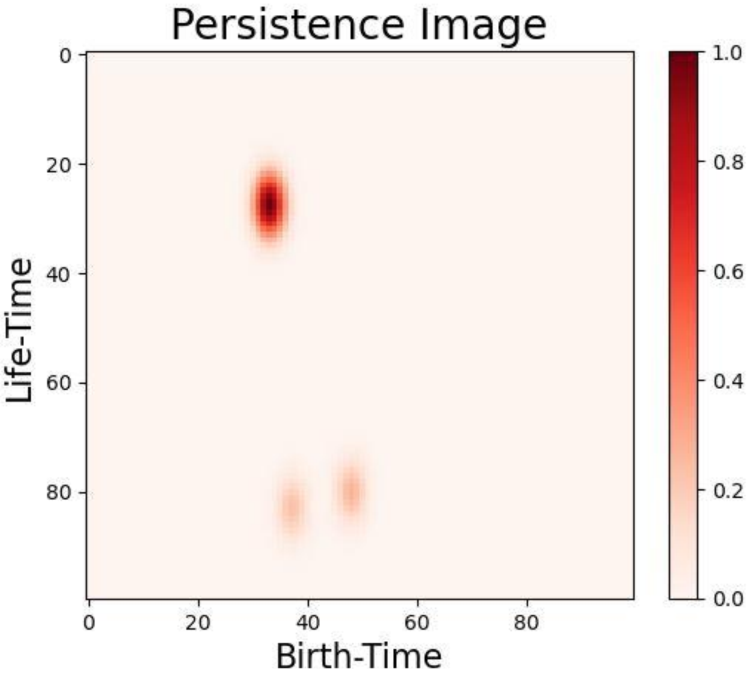
Persistence Images



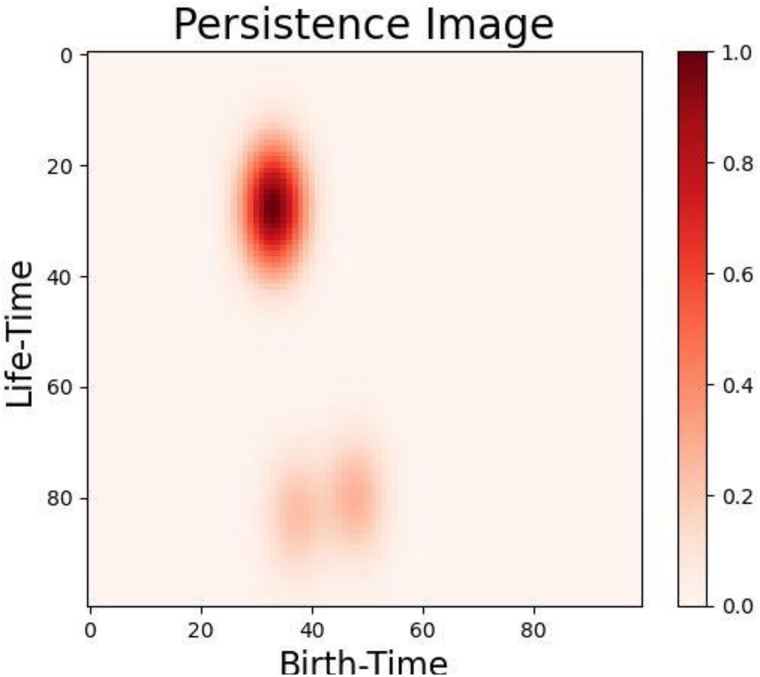


[6] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier. Persistence images: A stable vector representation of persistent homology. The Journal of Machine Learning Research, 18(1):218–252, 2017.

Effect of Spread

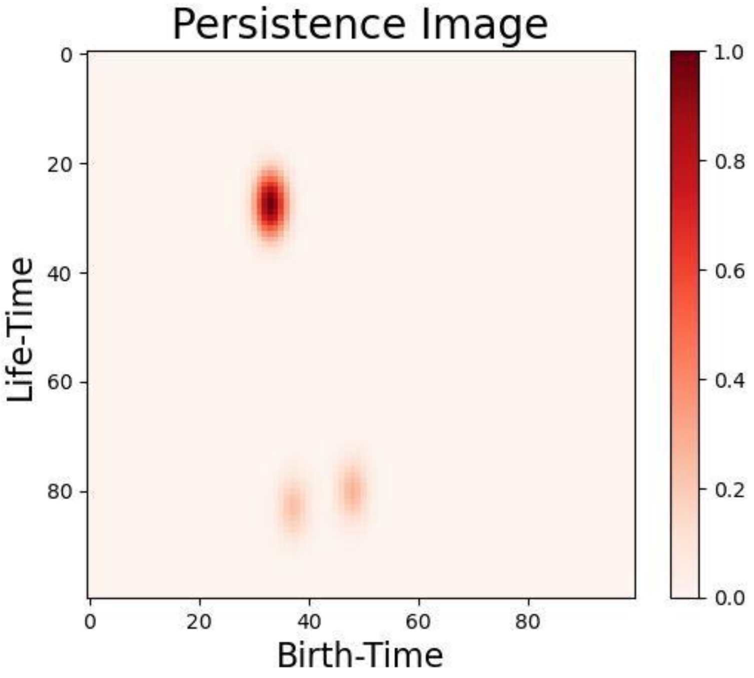


Spread = 0.1

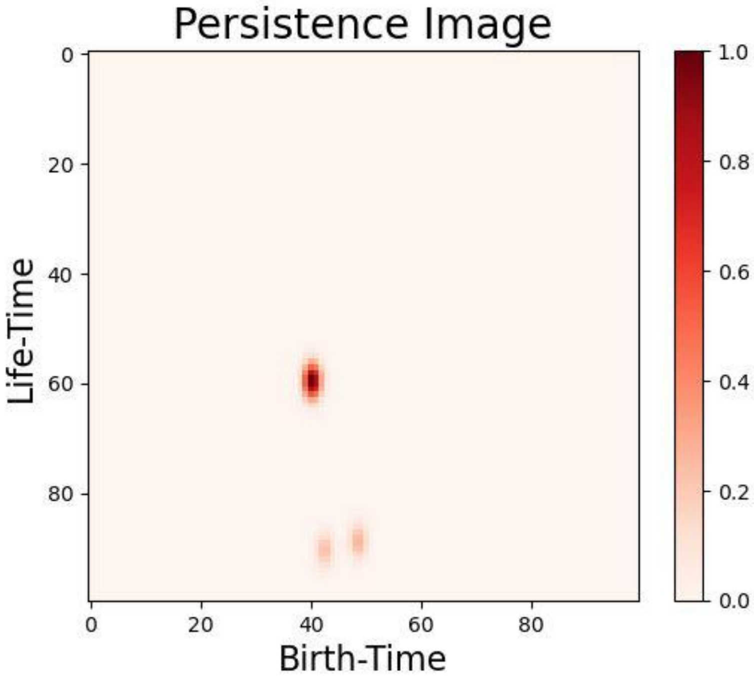


Spread = 0.2

Effect of Range

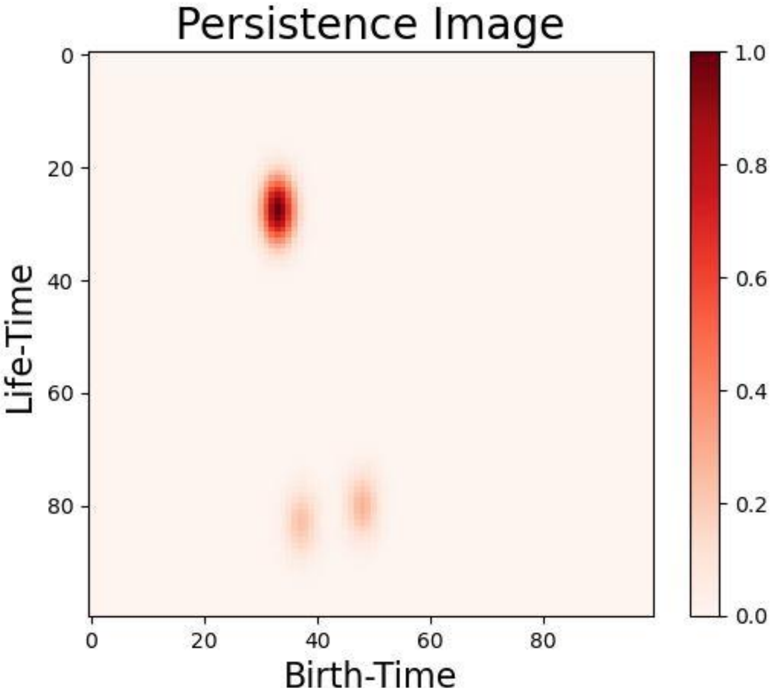


<u>Range = [-2.5, 2.5]</u>

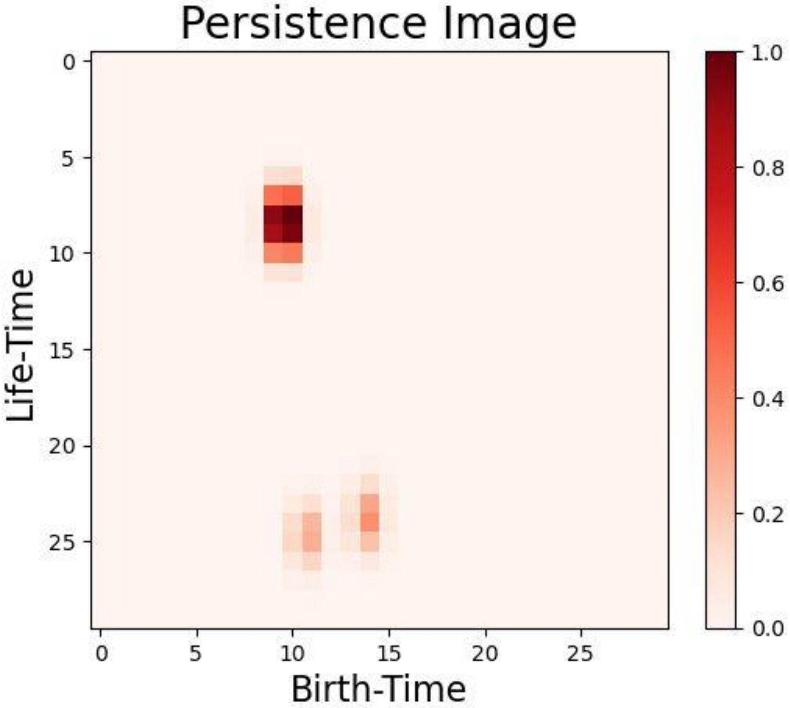


<u>Range = [-4.5, 4.5]</u>

Effect of Pixel Value



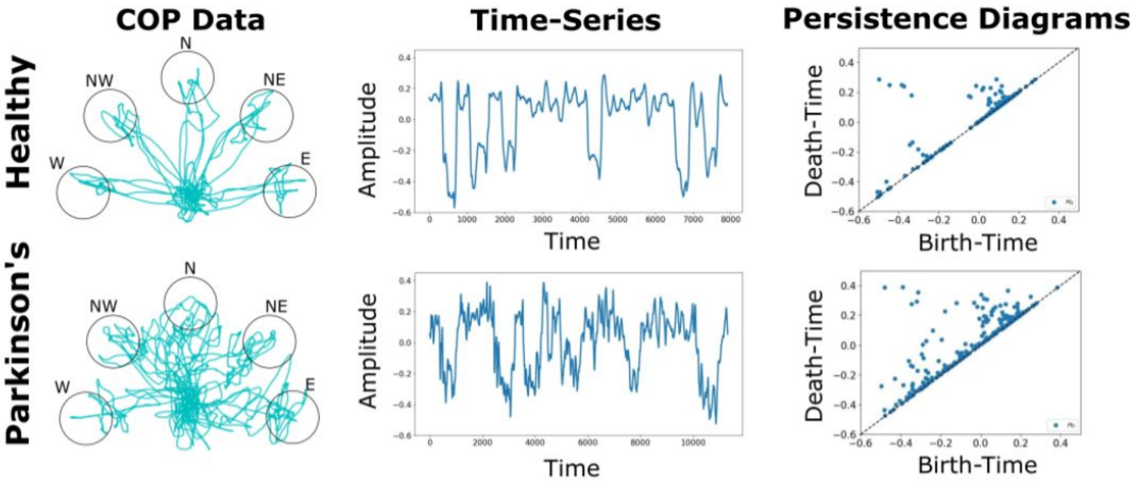
Pixel Grid: 100x100



Pixel Grid: 30x30

Application

Processing Flow



Round-Robin Leave-One-Subject-Out Cross Validation Protocol with Support Vector Machines

Binary Classification

- Parkinsons vs. Healthy

Three-Class Classification

- Parkinsons vs. Young vs. Elderly

Regression

- Parkinsons Patients Given Clinically Assigned UPDRS Scores
- Healthy Patients Given UPDRS Score of 0

Experimental Results

Method	3-Class Classification (%)	Regression	
		Correlation	P-Value
Peak Velocity Index	53.01	0.8153	$2.8227e^{-15}$
LLE	47.37	0.6449	$2.6707e^{-08}$
A3 [1]	60.53	0.7518	$4.4376e^{-12}$
D1 [1]	70.30	0.8479	$9.2763e^{-18}$
D2 [1]	71.43	0.9006	$1.1847e^{-22}$
D3 [1]	68.42	0.8479	$1.2715e^{-17}$
D4 [1]	65.79	0.8509	$7.2852e^{-18}$
A3M [7]	65.41	0.8528	$5.2806e^{-18}$
D1M [7]	73.31	0.8649	$5.1880e^{-19}$
D2M [7]	71.05	0.8639	$6.3778e^{-19}$
D3M [7]	71.80	0.8598	$1.4184e^{-18}$
D4M [7]	68.05	0.8680	$2.7529e^{-19}$
Persistence Images	**73.68**	**0.8493**	$\mathbf{9.85e^{-18}}$

Binary Classification: **98.87%**

[1] A. Som, N. Krishnamurthi, V. Venkataraman, and P. Turaga. Attractor shape descriptors for balance impairment assessment in parkinson's disease. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 3096–3100. IEEE, 2016
[7] A. Som, N. Krishnamurthi, V. Venkataraman, K. N. Ramamurthy, and P. Turaga. Multiscale evolution of attractor-shape descriptors for assessing parkinson's disease severity. In 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pages 938–942. IEEE, 2017.

Thank You

For any questions, please do not hesitate to email us at:

- Afra Nawar: anawar@asu.edu
- Farhan Rahman: fnrahman@asu.edu
- Anirudh Som: asom2@asu.edu

Also refer to our GitHub repository with sample code here: https://github.com/itsmeafra/Sublevel-Set-TDA