Topological Data Analysis of Diabetic Retinopathy and Macular Edema Images

Sruthi Suresh

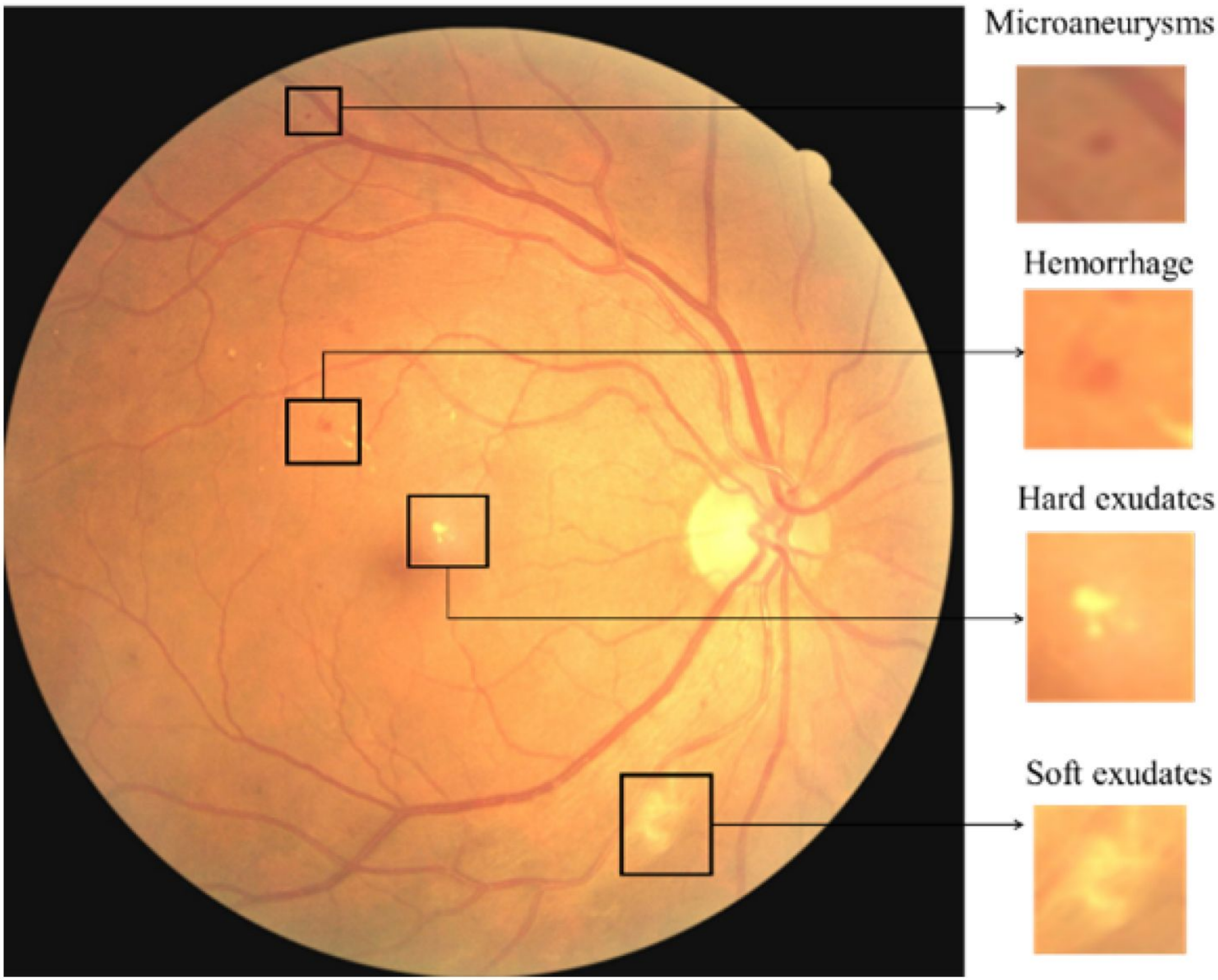
Math 4v91: Topological Data Analysis

December 8th, 2022

The Problem

- Diabetic Retinopathy (DR) is one of the leading causes of preventable blindness in the adult working population [1].
  - As the most common complication of Diabetes Mellitus (DM), estimated to affect 463 million people in 2019, it poses a pressing global health concern [2].
- Diabetic Macular Edema (DME) is the most common cause of vision loss from DR [3].

Signs



Fundus Photography

- Fundus photography captures the macula and optic nerve of the eye and is easily used.
- Provides a clear visualization of morphological changes including exudates.
- Allows for monitoring and diagnosis of the disease.

Motivation

- With the increase in patients who have diabetes and a lack of concurrent increase in retinal specialists or fundus image graders, it is important to explore automated options for DR and DME detection.

Retinal Fundus Images

Used the MESSIDOR Dataset [4].

- Consists of 1200 eye fundus color images acquired by 3 ophthalmologic departments (800 x 800).
- Medical experts diagnosed each image with:
  - Diabetic Retinopathy Grade
    - 0 (Normal): ($\mu A = 0$) AND ($H = 0$)
    - 2: (($5 < \mu A < 15$) OR ($0 < H < 5$)) AND ($NV = 0$)
    - 3: ($\mu A >= 15$) OR ($H >=5$) OR ($NV = 1$)

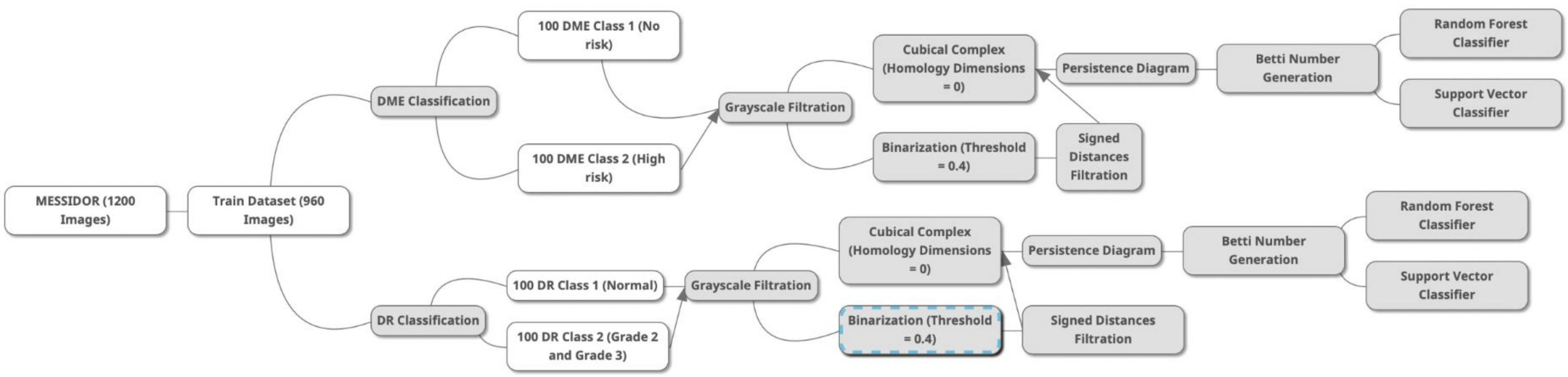
  μA: number of microaneurysms
  H: number of hemorrhages
  NV = 1: neovascularization
  NV = 0: no neovascularization

  - Risk of Macular Edema
    - 0 (No risk): No visible hard exudate
    - 2 (High risk): Shortest distance between macula and hard exudates <= one papilla diameter

Purpose

- Can a TDA-based approach classify with high accuracy:
  - Normal eyes and eyes with a high grade of retinopathy?
  - Eyes with no visible exudate (low risk of DME) and eyes with a high risk of DME?
- Additionally, can we understand how this would appear with clinical deployment?

Planned Overview



Pre-processing

1. Images were downloaded using the Kaggle API.
2. Sorted images into folders by DR classification and DME classification using train.csv column values as reference.
   a. DR Classification: Binary (Class 1 (DR = 1) and Class 2 (DR = 2 and 3))
   b. DME Classification: Binary (Class 1 and Class 2)
3. Randomly selected files using random.sample to build two folders of 200 images each for training purposes and then converted to grayscale.
   a. DR Classification: Binary (100 Class 1 and 100 Class 2 (50 of each subset))
   b. DME Classification: Binary (100 Class 1 and 100 Class 2)

Feature Extraction

1. Images were, once converted to grayscale
2. Used grayscale filtration and cubical complexes used to generate persistence diagrams and Betti-0.

Classification using ML

1. Outputted into dataframes with columns as X and Y as labels as directed in class.
2. Trained Random Forest Classifier (70%/30%) using:
   a. Betti Numbers of Dimension 0 (Grayscale Filtration)
3. Trained Support Vector Classifier (70%/30%) using:
   a. Betti Numbers of Dimension 0 (Grayscale Filtration)

3. Obtained the following metrics:

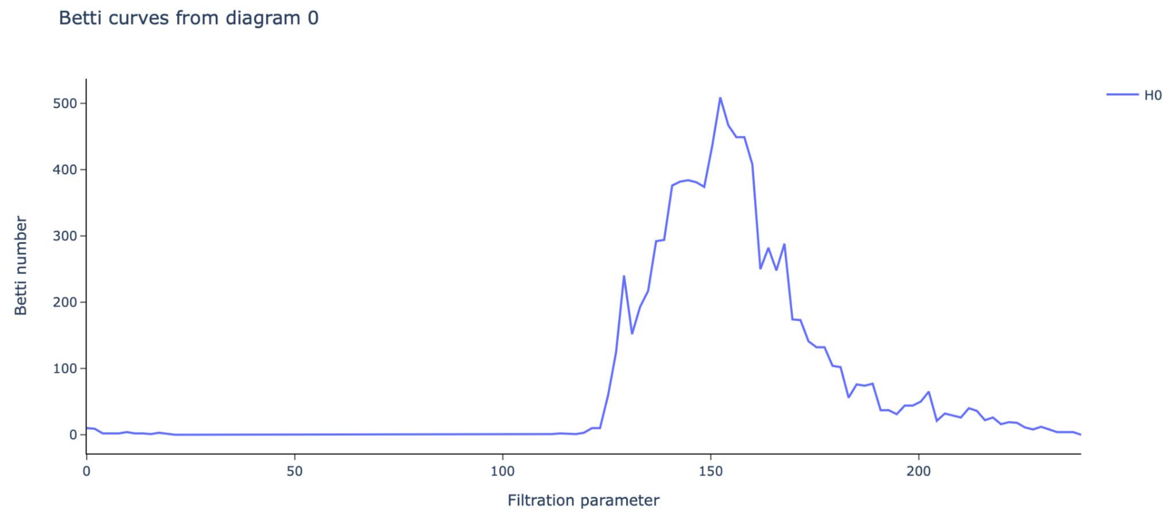
   a. Accuracy

   b. Confusion Matrix

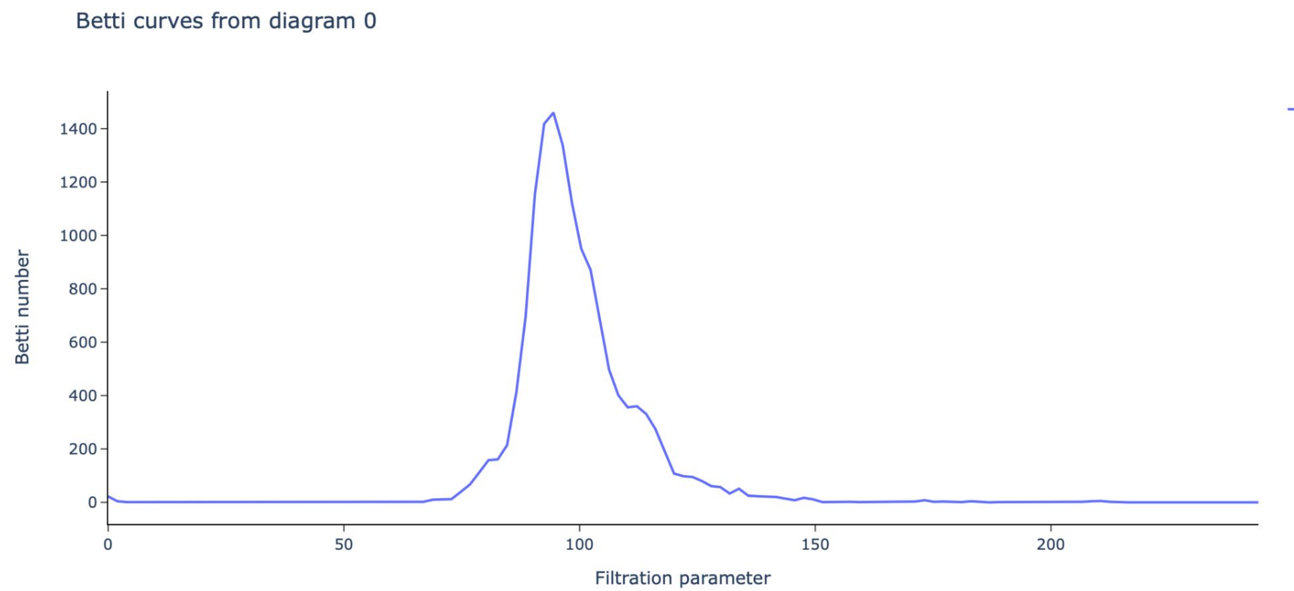
   C. AUC

DME (Grayscale), 125 Bins

No Risk:

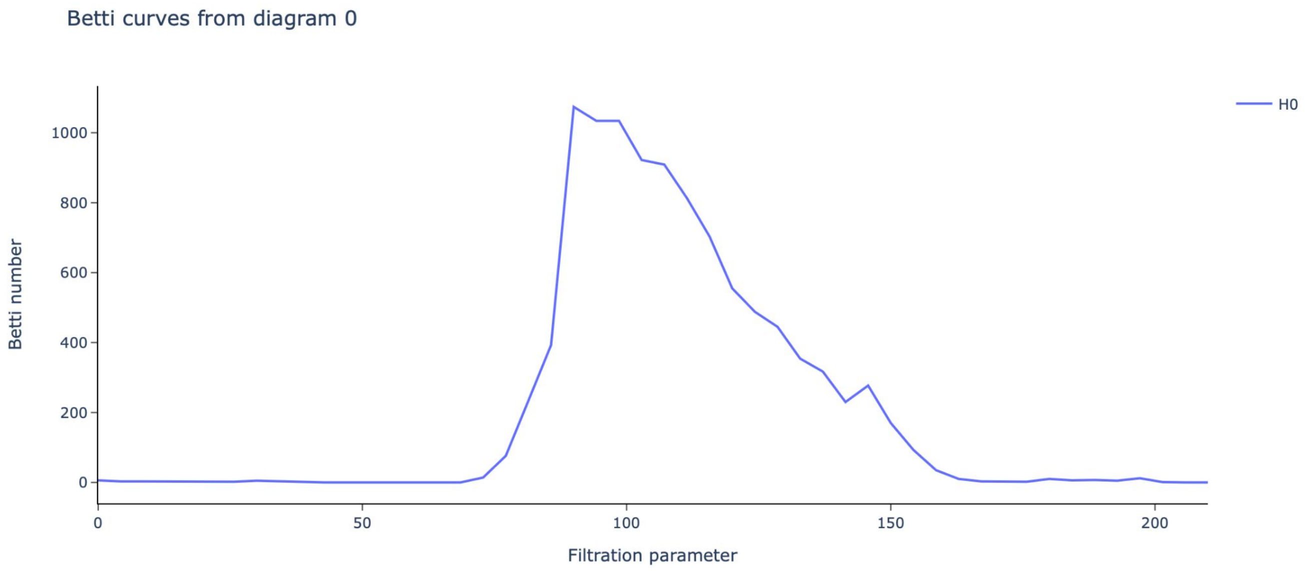


High Risk:

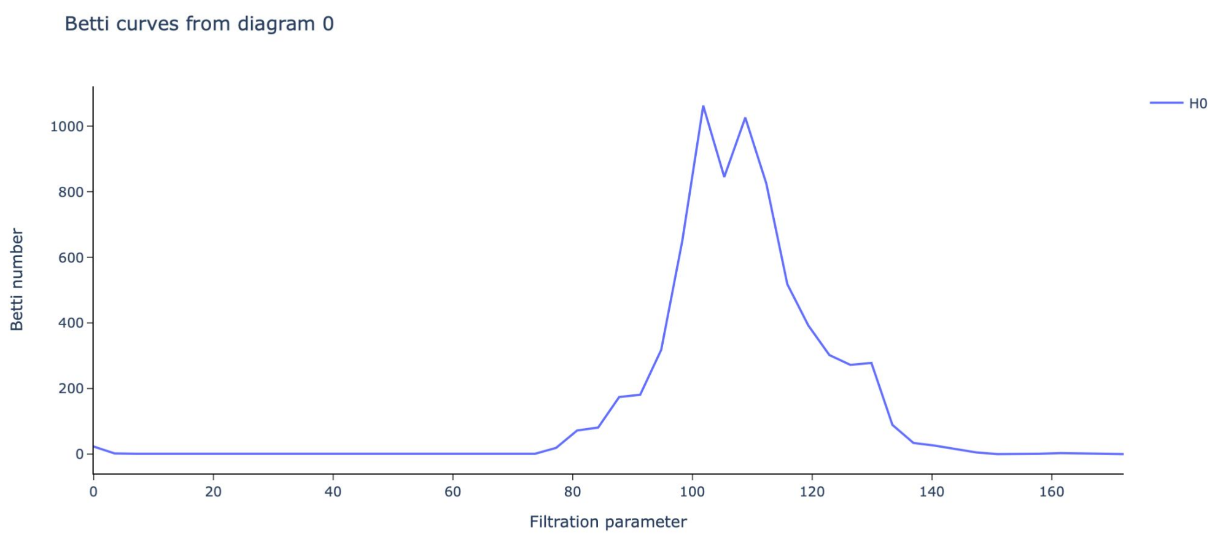


DR (Grayscale), 50 Bins

Normal:



Grade 2+:



DR Classification (Grayscale)

Model	Accuracy	AUC	Confusion Matrix
Random Forest	0.98	0.98	[28 1] [0 31]
Support Vector Classifier (Linear)	0.90	0.90	[27 2] [4 27]

Clinical Implementation

- Streamlit.io used for example of how to implement in clinical practice.
- Show Demo.

DME Classification (Grayscale)

Model	**Accuracy**	**AUC**	**Confusion Matrix**
Random Forest	0.58	0.59	[19 16] [9 16]
Support Vector Classifier	0.55	0.56	[18 17] [10 15]

Comparison (Selected)

Paper	Problem	AUC
Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs	Diabetic Retinopathy (Binary)	0.99
Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs	Diabetic Retinopathy (Binary)	0.853
Diabetic macular edema grading in retinalimages using vector quantization and semi-supervised learning	Edema (Three Classes)	0.975

Areas of Improvement

- Build SciKit Pipeline.
- Better Accuracy and AUC
- Explore parameter tuning for better results.
- Try other filtration options.
- Try different ML Options.
- Increase dataset size.

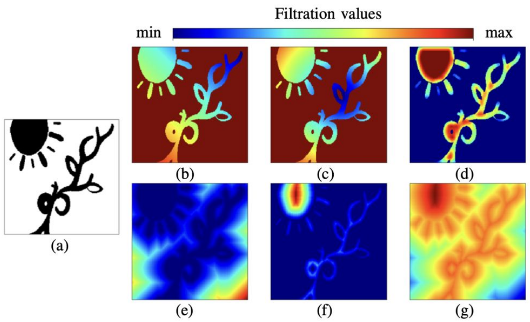


Fig. 5. Grayscale images obtained from the different types of filtrations. For visibility reasons, we choose a colored map to represent grayscale values. The original binary image we use to illustrate our filtrations is shown in (a) and also gives a binary filtration. The height filtration with vector $(-1, -1)$ is shown in (b), the radial filtration with center in the blue leaf in (c), the density filtration in (d), the dilation filtration in (e), the erosion filtration in (f) and the signed distance filtration in (g).

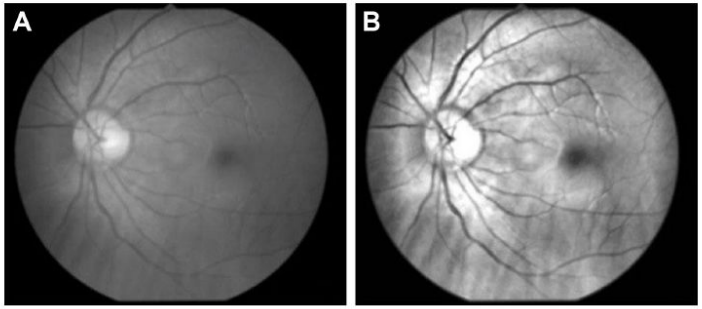


Figure 2 The result of the pre-processing method: (**A**) original image and (**B**) result image.

Works Cited

[1] Saeedi, Pouya et al. "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition." *Diabetes research and clinical practice* vol. 157 (2019): 107843. doi:10.1016/j.diabres.2019.107843

[2] Khalil, H. "Diabetes microvascular complications-A clinical update." *Diabetes & metabolic syndrome* vol. 11 Suppl 1 (2017): S133-S139. doi:10.1016/j.dsx.2016.12.022

[3] Mitchell P, Annemans L, Gallagher M, et al. Cost-effectiveness of ranibizumab in treatment of diabetic macular oedema (DME) causing visual impairment: evidence from the RESTORE trial.
[4] https://www.adcis.net/en/third-party/messidor/

Logo Image Here