

Article

Topological Data Analysis for Eye Fundus Image Quality Assessment

Gener José Avilés-Rodríguez ^{1,*}, Juan Iván Nieto-Hipólito ^{1,*}, María de los Ángeles Cosío-León ², Gerardo Salvador Romo-Cárdenas ¹, Juan de Dios Sánchez-López ¹, Patricia Radilla-Chávez ³ and Mabel Vázquez-Briseño ¹

¹ Facultad de Ingeniería Arquitectura y Diseño, Universidad Autónoma de Baja California, Carretera Transpeninsular Ensenada-Tijuana #3917, Playitas, Ensenada 22860, Mexico; romo.gerardo@uabc.edu.mx (G.S.R.-C.); jddios@uabc.edu.mx (J.d.D.S.-L.); mabel.vazquez@uabc.edu.mx (M.V.-B.)

² Dirección de Investigación, Innovación y Posgrado, Universidad Politécnica de Pachuca, Carretera Ciudad Sahagún-Pachuca Km. 20, Ex-Hacienda de Santa Bárbara, Hidalgo 43830, Mexico; ma.cosio.leon@upp.edu.mx

³ Escuela de Ciencias de la Salud, Universidad Autónoma de Baja California, Carretera Transpeninsular S/N, Valle Dorado, Ensenada 22890, Mexico; patyradilla@uabc.edu.mx

* Correspondence: gener.aviles@uabc.edu.mx (G.J.A.-R.); jnieto@uabc.edu.mx (J.I.N.-H.)

Abstract: The objective of this work is to perform image quality assessment (IQA) of eye fundus images in the context of digital fundoscopy with topological data analysis (TDA) and machine learning methods. Eye health remains inaccessible for a large amount of the global population. Digital tools that automatize the eye exam could be used to address this issue. IQA is a fundamental step in digital fundoscopy for clinical applications; it is one of the first steps in the preprocessing stages of computer-aided diagnosis (CAD) systems using eye fundus images. Images from the EyePACS dataset were used, and quality labels from previous works in the literature were selected. Cubical complexes were used to represent the images; the grayscale version was, then, used to calculate a persistent homology on the simplex and represented with persistence diagrams. Then, 30 vectorized topological descriptors were calculated from each image and used as input to a classification algorithm. Six different algorithms were tested for this study (SVM, decision tree, k-NN, random forest, logistic regression (LoGit), MLP). LoGit was selected and used for the classification of all images, given the low computational cost it carries. Performance results on the validation subset showed a global accuracy of 0.932, precision of 0.912 for label “quality” and 0.952 for label “no quality”, recall of 0.932 for label “quality” and 0.912 for label “no quality”, AUC of 0.980, F1 score of 0.932, and a Matthews correlation coefficient of 0.864. This work offers evidence for the use of topological methods for the process of quality assessment of eye fundus images, where a relatively small vector of characteristics (30 in this case) can enclose enough information for an algorithm to yield classification results useful in the clinical settings of a digital fundoscopy pipeline for CAD.

Keywords: persistent homology; eye fundus images; topological data analysis; image quality assessment; computational ophthalmology



Citation: Avilés-Rodríguez, G.J.; Nieto-Hipólito, J.I.; Cosío-León, M.d.l.Á.; Romo-Cárdenas, G.S.; Sánchez-López, J.d.D.; Radilla-Chávez, P.; Vázquez-Briseño, M. Topological Data Analysis for Eye Fundus Image Quality Assessment. *Diagnostics* **2021**, *11*, 1322. <https://doi.org/10.3390/diagnostics11081322>

Academic Editor: Takashi Kojima

Received: 3 June 2021

Accepted: 16 July 2021

Published: 23 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Public Health Dimension

Eye health has a profoundly multidimensional effect in overall health, economics, and social development for populations around the world [1]. Globally, there are more than 250 million people with vision impairment and over a billion with near-vision impairment [2]. It is projected that over the next 30 years, the amount of people affected by these issues will triple, reaching around 700 million, mostly due to the aging and growth of the population [2]. Even more, 90% of this loss occurs in low-income and middle-income

countries leading to a substantial economic burden with a global annual estimate of over USD 3 trillion [3].

Even though diagnostic and therapeutic strategies are available for the many causes of sight loss, for a significant amount of the global population at risk, they remain inaccessible mainly due to the lack of local eye care services and a considerable shortage and inefficient distribution of appropriately trained personnel [4].

To address these challenges, with enough evidence and with potential to scale, these strategies can focus on capacity building of clinical personnel and the use of technology to empower human resources [5]. Computational approaches in digital image analysis have been proposed as a strategy to strengthen and complement eye health teams, facilitating accessibility to health services for medically underserved populations, one of the areas in which computational tools have proved to be useful is that of digital image processing as computer-aided diagnosis (CAD) systems [6].

A fundamental aspect in the evaluation of eye health is the assessment of eye fundus through ocular fundoscopy [7].

This can be achieved by direct observation of the eye fundus through an ophthalmoscope or through eye fundus images. As the methodology to evaluate a fundoscopy is not universal, it is strongly recommended that a systematic and organized approach be taken to this approach in the clinical practice [8]. This will allow this approach to be adapted and reproduced with a computational system that mimics the evaluations a clinician would perform during a regular consultation.

Nonetheless, state-of-the-art algorithms used for CAD systems in eye fundus images [9–11] tend to need large volumes of images in their training stages in order for them to achieve acceptable performance indicators, as well as accurately labeled images [12]. Therefore, the exploration of approaches to produce robust results with a relatively smaller volume of training images is important, given the current tendency in the field towards the use of deep learning approaches that require significantly large annotated datasets, which is a current challenge in medicine [13]. This will allow CAD tools to adjust to a clinical environment more rapidly, thus, facilitating expected outcomes.

1.2. Fundus Image Analysis

Fundus image analysis can be understood as the process of obtaining a digital image of the eye fundus and the analytical pipeline required to generate a CAD tool to support physicians in their clinical practice [14,15]. Eye fundus imaging is the most established technique of retinal imaging; Figure 1 shows a summary of the components considered for this process.

As proposed by Abramoff [14], image quality assessment (IQA) is considered the first step in automated analysis techniques of eye fundus images. Development of IQA algorithms depends on the clinical application of the overall analysis [16–18] and can be classified in three general groups according to the techniques in which they achieve their objective:

1. Image quality parameters.

These algorithms are based on parameters like clarity, focus, contrast, and illumination. They generally have low computational complexity and are preferred when using mobile or low powered devices, some examples can be seen in [19,20] and [21].

2. Based on segmentation.

Techniques such as image structure clustering [22] or segmentation maps and feature analysis fall on this category [23].

3. Deep learning.

The most recent approach in IQA, done with convolutional neural networks (CNNs) [9] and human-in-the-loop approaches combined with CNNs [24].

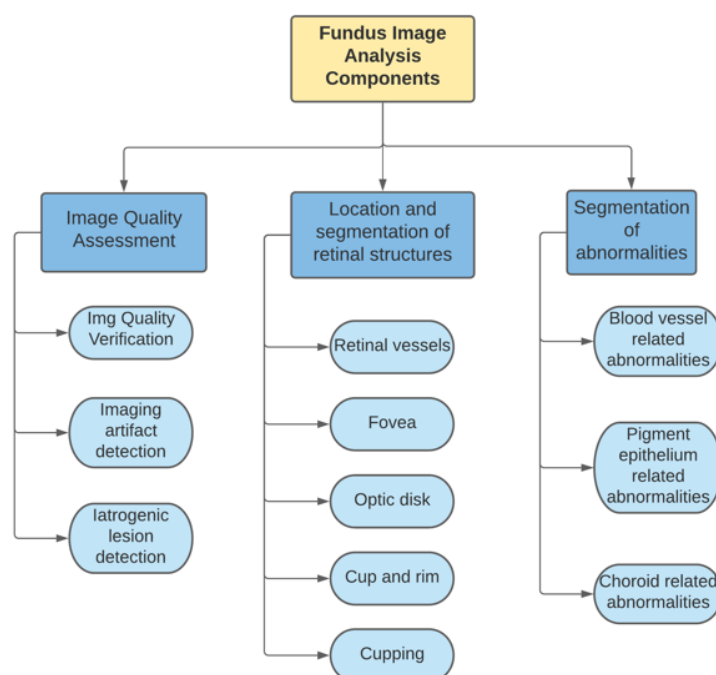


Figure 1. Fundus image analysis components as proposed by [14].

1.3. Topological Data Analysis

Topology mainly focuses on the analysis of the characteristics preserved under continuous deformation in geometric objects [25] (for example, noise or missing data represented in point clouds). In recent years, there has been a significant rise in computational approaches to topological concepts for the analysis of data [26–28]; this area is known as topological data analysis (TDA).

For the analysis of medical data, TDA becomes an attractive tool given the robustness of the methods to missing and noisy data, which are common in this domain.

In particular, for the case of images, they can be interpreted into mathematical objects called cubical simplicial complexes and then analyzed with topological methods to obtain characteristics to be used later in the analytical pipeline (Section 3.2). In this work, we focus on the interpretation of eye fundus images as cubical simplicial complexes and the obtention of homological groups (Betti numbers, $\beta_{\{1,2\}}$) to be used in the process of selecting optimal threshold values for binarizing a grayscale version of the image to obtain masks for background segmentation.

2. Materials and Methods

Eye fundus images used for this project were taken from the EyePACS project [29]. For the particular task of image quality assessment, we used labels proposed in [30], as presented by [31]. We modeled the phenomenon as a binary classification between images with quality (good) and those without quality (bad) and also adjusted for class imbalance by randomly selecting subsets by label of 2000 images each. Figure 2 shows a sample of images by label.

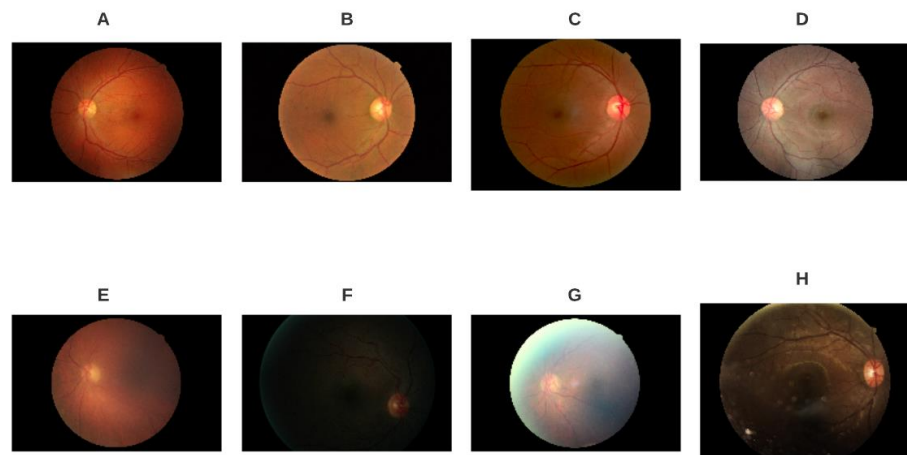


Figure 2. Examples of EyePACS images selected for the study with label good quality (A–D) and bad quality (E–H). Notice the variability within labels given by color, illumination, blurring, and anatomical landmarks.

3. Methods

3.1. Topological Interpretation of Digital Images

3.1.1. Cubical Complexes for the Representation of Digital Medical Images

Considering the need to use mathematical objects to represent 2D digital images, cubical complexes arise as a natural representation for this domain [32] given that pixels can be represented by 2-cubes and voxels by 3-cubes, as shown in Figure 3. This approach has already been used by [33] to analyze images of blood vessels by interpreting segmented images from magnetic resonance imaging as cubical simplexes and calculating their homological groups.

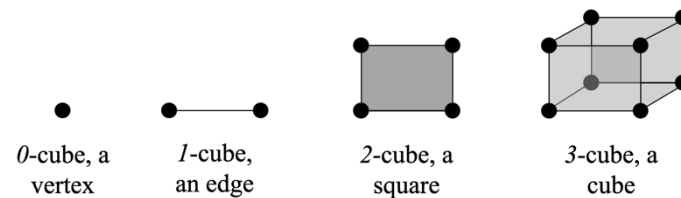


Figure 3. Visual summary of the general process for topology-regulated background extraction of eye fundus digital images.

A cubical complex is a combinatorial structure used in topological analysis where a 0-cube is called a vertex, a 1-cube, an edge, a 2-cube a square, and a 3-cube, a cube. This concept is built on the theory of elementary intervals and cubes; an in-depth approach to the theory of cubical complexes is presented by [34].

In a cubical complex (K) , multidimensional cubes (σ^d) play the role of simplices, where a finite cubical complex in \mathbb{R}^d is a union of cubes aligned on the grid \mathbb{Z}^d satisfying specific conditions as with simplicial complexes.

A d -dimensional digital image can be considered as a discrete map $\mathcal{I} : I \subseteq \mathbb{Z}^d \rightarrow \mathbb{R}$. In this context, an element $v \in I$ is called a pixel when $d = 2$ and a voxel when $d = 3$. $\mathcal{I}(\square)$ is the correspondent intensity or grayscale value. For the case of a binary image, the discrete map is $\mathcal{B} : I \subseteq \mathbb{Z}^d \rightarrow \{0, 1\}$.

For the purpose of this work, we choose to represent images as cubical complexes where a pixel is represented by a d -cube, and all its faces (adjacent lower-dimensional cubes) are added. A function on the resulting cubical complex K by extending the values of the pixels to all the cubes σ in K as the next formula shows:

$$\mathcal{I}'(\sigma) := \min_{\sigma \text{ face of } \tau} \mathcal{I}(\tau) \tag{1}$$

Therefore, a grayscale image comes with a natural filtration embedded in the grayscale values of its pixels. Let K be the cubical complex built from the image I . Then, let:

$$K_i := \{\sigma \in K \mid \mathcal{I}'(\sigma) \leq i\} \tag{2}$$

for the i -th sublevel set of K . The set $\{K_i\}_{i \in \mathcal{I}}$ defines a filtration of cubical complexes, indexed by the value of the grayscale function \mathcal{I} .

All images in the EyePACS dataset are in the 8-bit format; therefore, the grayscale versions have a depth of luminescence of 256 values (0 for absolute black, 255 for absolute white, and shades of gray for the corresponding values in between them). According to the previous information and following Formula (2), a natural filtration with 256 sublevels can be extracted from these images once they are interpreted as 2-cubical simplices given that $i = \{0, 1, 2, \dots, 256\}$.

Figure 4C exemplifies the issue of a pixelated version of Figure 4B where pixels are visible as cells in a 2D grid (the image matrix) and naturally interpretable as 2-cubes (see Figure 2), each one of the four sides of a pixel (lines delimiting each cell in the image matrix) can be interpreted as a 1-cube (see Figure 2). Following a combinatorial approach, we obtain a mathematical object representing the digital grayscale image on which topological data analysis can be performed, namely the 2D cubical complex.

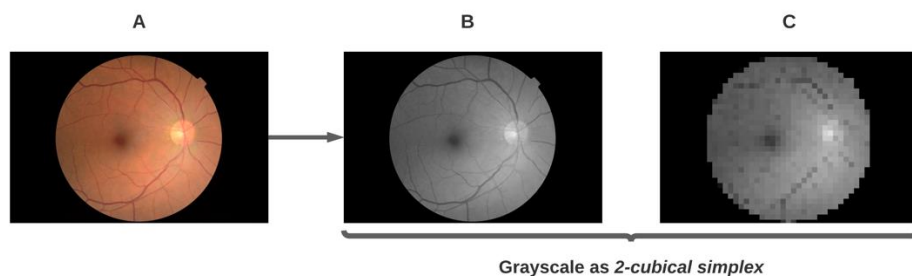


Figure 4. Example of the process of transforming a color eye fundus image from the EyePACS dataset (A) to the grayscale version (B) and a simplification to show how pixels in a grayscale image fulfill the definition of a cubical complex in 2 dimensions (C).

3.1.2. Cubical Filtrations

The steps to obtain a filtration from a grayscale image are then:

$$\text{Image} \rightarrow \text{Cubical complex} \rightarrow \text{Sublevel sets} \rightarrow \text{Filtration.}$$

Figure 5 shows a filtration of a cubical complex induced by the grayscale values contained in the image matrix and the corresponding barcode; notice that the barcode only has one element per homological group (0 and 1 dimensional topological characteristics). This information is reported in the form of Betti numbers (β_i); where β_0 = the number of 0-dimensional topological characteristics or connected components (connected complete black pixels for the use case of this work) at a specific sublevel filtration, β_1 = the number of 1-dimensional topological characteristics or holes inside connected components (complete white pixels inside components made of black pixels). This analysis is called persistent homology (PH) [35]. The figure presents an oversimplified example. Real data will typically have multiple bars per homological group in a filtration as seen in Figure 6.

In the example shown in Figure 5, only one connected component appears at filtration time 1 and survives until the end, as all the new pixels are always connected to the previous ones. A 1-dimensional homological group (H_1) appears at filtration time 2 and is filled in at filtration time 4. Figure 6 presents barcodes for the grayscale eye fundus image shown in Figure 4B. As it can be appreciated, there is a significant quantity of barcodes with birth and death values close to each other (short bars) in both H_0 and H_1 . The structures represented by the short bars in the barcode are considered topological noise and are usually discarded. Longer bars represent structures, connected components for H_0 or holes for H_1 , that persist

through a significant amount of filtration sublevels and could hold information relevant for processes in the clinical diagnosis.

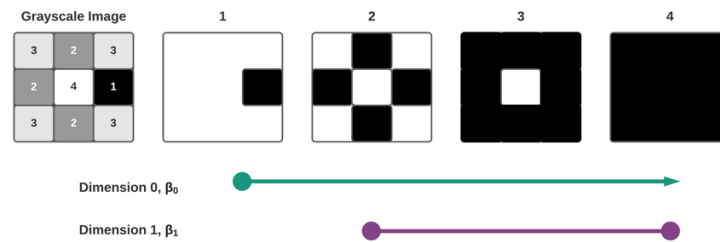


Figure 5. Induced filtration of a 3×3 grayscale image matrix and the correspondent barcode for β_0 and β_1 . Using a luminescence depth of 4 elements, $L = \{1, 2, 3, 4\}$, to simplify the example.

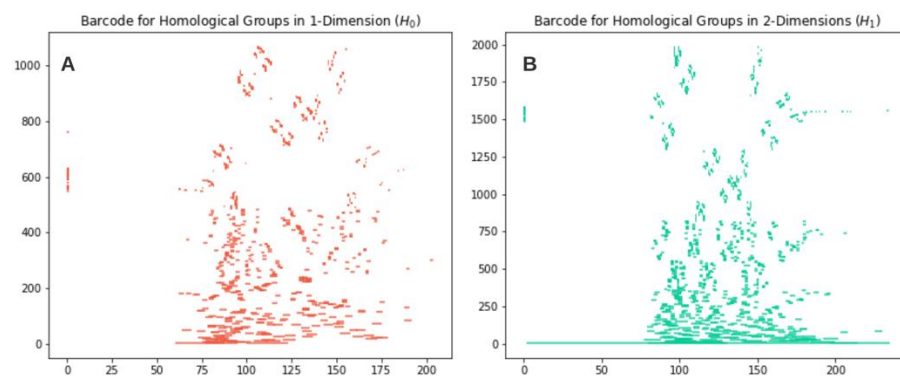


Figure 6. Corresponding bar codes for 0 and 1-dimensional homological groups (A,B) of Figure 4B. Displayed on the x axis are the filtrations sublevels (since it is an 8-bit image there are 256 sublevels available). The y axis shows a count of structures per sublevel in the form of bar codes. Each barcode starts and finishes at the birth and dead sublevel value for the structure it represents.

Figure 6A,B shows at least one bar that persists longer than the rest for each dimension (H_0 and H_1). Figure 7 explores in a visualization some of the sublevel sets (grayscale values) corresponding to each bar per dimension.

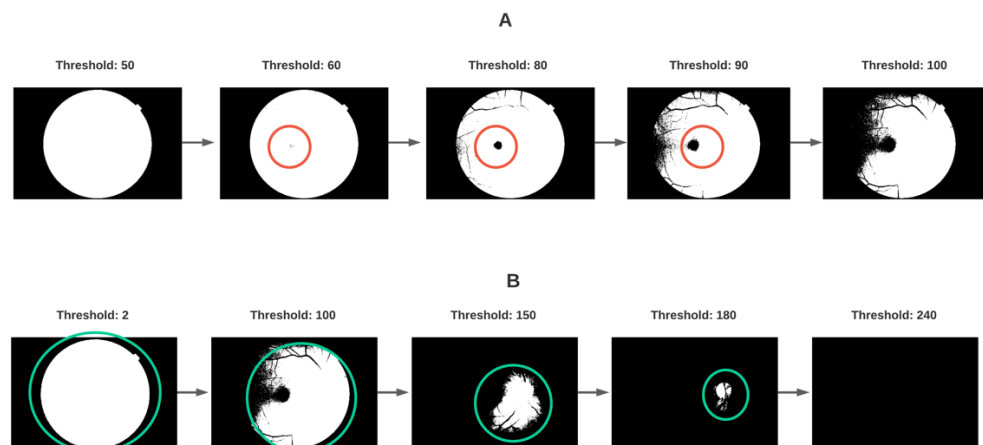


Figure 7. Visualization of binarized images from Figure 4B thresholded at grayscale values contained in the topological elements persisting for the largest amount of sublevel sets in the filtration represented by barcodes in Figure 6. Done for the bar in H_0 (A) and H_1 (B).

By binarizing Figure 4B using threshold values contained in the longest bar of Figure 6A, we can appreciate the resulting series in Figure 7A; a connected component of clinical relevance appears at around a grayscale value of 60 and persists until a value of 90

before merging into a larger unified component with the background at around a grayscale value of 100. Interestingly this spot, connected component, or 0-dimensional homological group is consistent with the macula and fovea as anatomical landmarks. It is here at this point that the topological results begin to connect with clinical meaning for the domain from which the data come, in this case eye fundus images and the diagnostic process in which they are used.

Figure 6B also shows a 1-dimensional homological group persisting significantly more than the rest in the filtration (from 2 to over 200). This means that there must be a recognizable hole (H_1 homological group) when binarizing the image using threshold values contained in the bar for such homological group. Figure 7B shows a series of binarized images using some of the values in the range of 2 to 240 and visually confirms the topological findings with a persistent hole, which first represents all the area of the camera sensor in which the information of the eye fundus image is contained and shrinks until the anatomical region where the optic disk would be expected to be. This is consistent with clinical practice where the area of the optic disk is typically regarded as the brightest in a healthy eye fundus evaluation, again, connecting 1-dimensional topological structures to clinical meaning.

3.2. Topological Indicators Derived from Digital Images

Once an eye fundus image has been interpreted as a 2D cubical complex and a corresponding PH analysis has been performed, as shown in Section 3.1. The results must be vectorized to provide an input that machine learning (ML) algorithms can take for classification purposes. For this work, ML algorithms are used to perform image quality assessment of the eye fundus images. This section presents details dealing with the process of vectorization of topological results represented by persistence diagrams.

Figure 8 shows a representation of the pipeline used in this work to perform topological data analysis (TDA) and obtain a vectorized representation to feed a ML classifier (in this case a logistic regression). Each eye fundus image was first converted from color to grayscale (Figure 4) and, then, entered the pipeline described in this section.

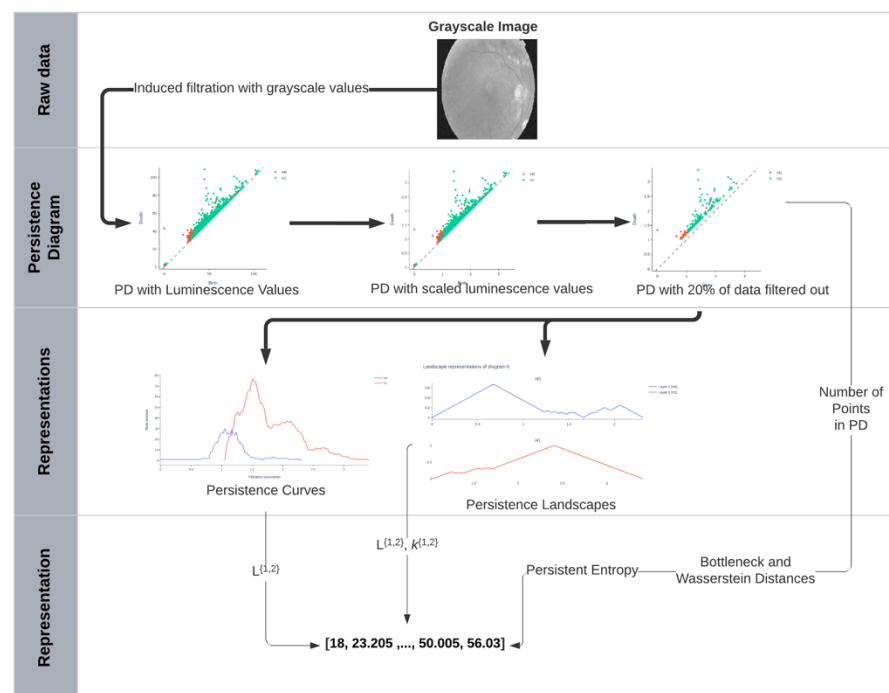


Figure 8. Representation of pipeline for extraction of topological descriptors of eye fundus images.

3.2.1. Persistence Diagrams

A persistence diagram (PD) is a visual representation of a set of points $\{(b, d) | b, d \in \mathbb{R}^2\}$ and $d > b$, where d = death of the topological feature and b = birth of the topological feature [36]. In order to be able to feed topological features represented in a persistent diagram to a machine learning algorithm they must be vectorized.

Figure 9 shows the corresponding PD for 0 and 1-dimensional topological characteristics obtained from a PH analysis of Figure 4B. Notice that the results are consistent with those observed in the code bars in Figure 6, only that it is easier to identify persistent characteristics.

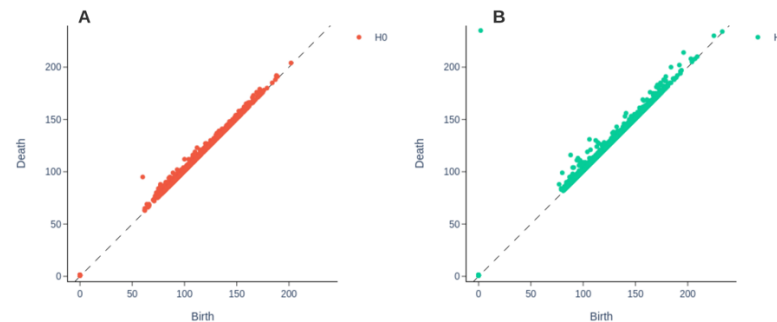


Figure 9. Persistence diagrams (PDs) corresponding to barcodes in Figure 6. Notice that the same information is presented in a more understandable manner, where the diagonal line represents the birth value and the points above the diagonal represent the death value of a given topological characteristic. PDs facilitate the identification of noise versus relevant topological characteristics.

The following paragraphs explain how results shown in a PD are vectorized for the analysis implemented in this work. From this point on, eye fundus images are represented by their corresponding PDs over which the rest of the analysis is performed.

3.2.2. Persistent Entropy of Persistence Diagrams

It is an intuitive measure of the entropy of the points in a persistence diagram. It results from extracting the Shannon entropy of the persistence (lifetime) of all cycles (topological structures) [37,38].

Let $D = \{(b_i, d_i)\}_{i \in I}$ be a persistence diagram with each $d_i < +\infty$, where b_i = the i -th topological birth point of the structure, and d_i = the i -th topological dead point of the structure. The persistence entropy of D is defined by:

$$PE(D) = \sum_{i=1}^n \frac{l_i}{L(B)} \log\left(\frac{l_i}{L(B)}\right) \tag{3}$$

where $L(B) := l_1 + \dots + l_n$ and $L(B) := l_1 + \dots + l_n$.

3.2.3. Bottleneck Distance

Let X and Y be two persistence diagrams. To define the distance between them, we consider bijections $\eta : X \rightarrow Y$ and record the least upper bound (*sup*) of the distances between corresponding points for each.

Measuring distance between points $x = (x_1, x_2)$ and $y = (y_1, y_2)$ with L_∞ -norm : $\|x - y\|_\infty = \max\{|x_1 - y_1|, |x_2 - y_2|\}$ and taking the greatest lower bound (*inf*) over all bijections, we get the bottleneck distance between diagrams [39], as shown in Equation (4).

$$W_\infty(X, Y) = \inf_{\eta: X \rightarrow Y} \sup_{x \in X} \|x - \eta(x)\|_\infty \tag{4}$$

A drawback of the bottleneck distance is that it is insensitive to details of the bijection beyond the furthest pair of corresponding points.

3.2.4. p -Wasserstein Distance

The p -Wasserstein distance between X and Y for any positive real number p , takes the sum of p -th powers of the L_∞ distances between corresponding points, again minimizing over all bijections, as shown in Equation (5).

$$W_p(X, Y) = \left[\inf_{\eta: X \rightarrow Y} \sum_{x \in X} \|x - \eta(x)\|_\infty^p \right]^{1/p} \quad (5)$$

It is also known as the Earth's movers distance, because intuitively, it can be interpreted as the minimum energy cost of moving and transforming a pile of dirt in the shape of one probability distribution to the shape of the other distribution. Therefore, the q -Wasserstein distance measures the similarity between two persistence diagrams using the sum of all edge lengths [40].

3.2.5. Persistence Landscape

The k -th persistence landscape of a barcode $\{(b_i, d_i)\}_{i=1}^n$ in the function $\lambda_k : \mathbb{R} \rightarrow [0, \infty)$ is the k -th largest value of $\{f_{(b_i, d_i)}(x)\}_{i=1}^n$, with:

$$f_{(b,d)}(x) = \begin{cases} 0 & \text{if } x \in (b, d) \\ x - b & \text{if } x \in \left(b, \frac{b+d}{2}\right) \\ -x + d & \text{if } x \in \left(\frac{b+d}{2}, d\right) \end{cases} \quad (6)$$

The parameter k is called the layer. In this work, we consider curves obtained when $k \in \{1, 2\}$ [41,42].

3.2.6. Betti Curves

For the purpose of this work, the Betti curve $B_n : I \rightarrow \mathbb{R}$ of a barcode $D = \{(b_i, d_i)\}_{i \in I}$ is the function that returns for each step $i \in I$, the number of bars (b_j, d_j) that contain i , as shown in Equation (7).

$$i \mapsto \#\{(b_j, d_j), i \in (b_j, d_j)\} \quad (7)$$

where $\#$ stands for cardinality.

3.2.7. Gaussian Kernel

By placing Gaussians of standard deviation σ over every point of the persistence diagram and a negative Gaussian of the same standard deviation in the mirror image of the points across the diagonal, the output of this operation is a real-valued function on \mathbb{R}^2 . For this work, we use $\sigma \in \{1.6, 3.2\}$ [43].

3.2.8. Number of Points in Persistence Diagram

This indicator refers to the number of off-diagonal points in a given persistence diagram, per homology dimension.

Given a persistence diagram consisting of birth–death–dimension triples $[b, d, q]$, subdiagrams corresponding to distinct homology dimensions are considered separately, and the respective numbers of off-diagonal points are counted and reported as the result.

3.3. Machine Learning Classifiers

Once the topological indicators are vectorized as mentioned in Section 3.2, the dataset is ready to be processed by standard machine learning classifiers.

For this work, we explored the following algorithms:

- Support vector machine;
- Classification tree;
- k-nearest neighbors;

- Random forest;
- Logistic regression;
- Multilayered perceptron.

For these initial explorations, the authors used Orange3 [44], selecting the default hyperparameters proposed by the software for each of the algorithms. Then, a subset of algorithms was selected based on their performance and evaluated in more depth in order to choose the classifier to use for this work.

3.4. Metrics for Evaluation of Performance of Classification Algorithms

Given a confusion matrix, as shown in Figure 10, in the context of a binary classification, the following indicators are identified:

- **True positives (TP):** entities classified by the algorithm as true to the label evaluated when the reference is also true.
- **True negatives (TN):** entities classified by the algorithm as true to the label evaluated when the reference is false.
- **False positives (FP):** entities classified by the algorithm as false to the label evaluated when the reference is also false, also known as Type I Error.
- **False negatives (FN):** entities classified by the algorithm as false to the label evaluated when the reference is true, also known as Type II Error.

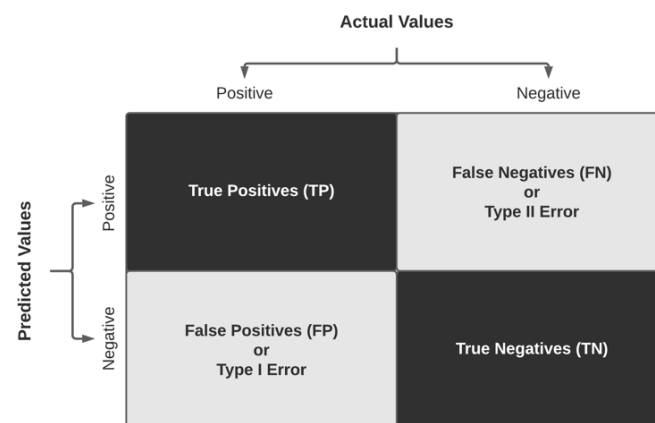


Figure 10. Visualization of a generic confusion matrix.

The following metrics are used to report on algorithm performance on this work [45]:

- Accuracy:

This metric answers the following question: overall, how often is our model correct? It is calculated using Equation (8):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

It does not work well with class imbalance, nor does it give detailed information about the application of the problem.

- Precision:

This metric answers the following question: what is the ratio between the true positives and all the positives? It is calculated using Equation (9):

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

This metric helps when the impact (defined by the domain of the data) of false positives is significantly high.

- Recall:

This metric answers the following question: what is the measure of our model correctly identifying true positives? It is calculated using Equation (10):

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

This metric helps when the impact (defined by the domain of the data) of false negatives is significantly high. It is also known as sensitivity or true positive rate (TPR).

- F1-score:

This metric is a combined measure of precision and recall. Therefore, a relatively high F1-score will mean that there are low false positives and low false negatives in the results generated by the classifier. It is calculated using Equation (11):

$$F1 = 2 \times \frac{\frac{TP}{TP+FP} \times \frac{TP}{TP+FN}}{\frac{TP}{TP+FP} + \frac{TP}{TP+FN}} = 2 \times \frac{precision \times recall}{precision + recall} \quad (11)$$

- Receiver-operating characteristic (ROC) curve:

This curve is considered a comprehensive performance measure and it is obtained by visualizing the true positive rate (TPR) versus the false positive rate (FPR), this last one is calculated following Equation (12):

$$FPR = 1 - Specificity = \frac{FP}{FP + TN} \quad (12)$$

- Area under the curve (AUC):

Once a ROC curve is generated, the area under such curve can also be used to measure the performance of the algorithm. The AUC shows the probability that a randomly classified positive sample becomes a higher score than a randomly classified negative sample. It can be expressed as shown in Equation (13):

$$AUC = P(X_{pos} > X_{neg}) \quad (13)$$

- Matthews correlation coefficient (MCC):

Typically used to evaluate the performance of classifiers when dealing with data with an unbalanced proportion of elements in each of the labels of the target variable [46], it is obtained by following Equation (14):

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (14)$$

4. Results

As shown in Section 3.2, each eye fundus image is represented by a vector of 30 topological descriptors and, then, given to a machine learning classification algorithm to predict image quality for clinical use.

Table 1 shows the resulting topological descriptors after the extraction and vectorization of the 30 topological indicators per image, as explained in Section 3.2.

Table 1. List of the 30 topological indicators calculated per image.

Variables 1–6	Variables 7–12	Variables 13–18	Variables 19–24	Variables 25–30
Persistence entropy β_0	2-Wasserstein distance β_0	Persistence landscape $L^2, k = 1, \beta_0$	Betti curve L^2, β_0	Gaussian kernel $L^2, \sigma = 1.6, \beta_0$
Persistence entropy β_1	2-Wasserstein distance β_1	Persistence landscape $L^2, k = 1, \beta_1$	Betti curve L^2, β_1	Gaussian kernel $L^2, \sigma = 1.6, \beta_1$
Bottleneck distance β_0	Persistence landscape $L^1, k = 1, \beta_0$	Persistence landscape $L^2, k = 2, \beta_0$	Gaussian kernel $L^1, \sigma = 1.6, \beta_0$	Gaussian kernel $L^2, \sigma = 3.2, \beta_0$
Bottleneck distance β_1	Persistence landscape $L^1, k = 1, \beta_1$	Persistence landscape $2, k = 2, \beta_1$	Gaussian kernel $L^1, \sigma = 1.6, \beta_1$	Gaussian kernel $L^2, \sigma = 3.2, \beta_1$
1-Wasserstein distance β_0	Persistence landscape $L^1, k = 2, \beta_0$	Betti curve L^1, β_0	Gaussian kernel $L^1, \sigma = 3.2, \beta_0$	Number of points in diagram β_0
1-Wasserstein distance β_1	Persistence landscape $L^1, k = 2, \beta_1$	Betti curve L^1, β_1	Gaussian kernel $L^1, \sigma = 3.2, \beta_1$	Number of points in diagram β_1

Preliminary explorations for the selection of a suitable algorithm were done on six classification algorithms (classifiers). Table 2 shows the performance metrics of the algorithms used in this phase. From these results, three algorithms were selected for the next steps in the process: SVM, MLP, and LoGit.

Table 2. Performance metrics from classification algorithms initially evaluated.

Model	AUC	CA	Precision	Recall	F1-Score
Support Vector Machine (SVM)	0.845	0.749	0.761	0.749	0.746
Decision Tree	0.870	0.894	0.895	0.894	0.894
k -Nearest Neighbors (k -NN)	0.941	0.898	0.900	0.898	0.898
Random Forest (RFC)	0.960	0.911	0.912	0.911	0.911
Logistic Regression (LoGit)	0.974	0.925	0.925	0.925	0.925
Multilayer Perceptron (MLP)	0.981	0.935	0.935	0.935	0.935

Where AUC = area under the curve and CA = classification accuracy.

In the following step, the authors performed a fine-tuning of hyperparameters of the three selected algorithms. Table 3 shows the resulting performance metrics after this process. Finally, the algorithm selected for the classification task was a LoGit because of the robust performance metrics it yielded and the relatively low computational cost when compared with both SVM and MLP.

Table 3. Performance metrics of fine-tuned classification algorithms.

Algorithm	Precision Training Set	Precision Testing Set	Recall Training Set	Recall Testing Set	F1-Score Training Set	F1-Score Testing Set
SVM	0.961	0.957	0.961	0.957	0.961	0.957
MLP	0.910	0.930	0.910	0.930	0.910	0.930
LoGit	0.989	0.987	0.989	0.987	0.989	0.987

Table 4 shows the hyperparameters and values used for the tuning process for LoGit. The optimal results were obtained when tolerance = 1×10^{-8} , C = 150,000, solver = liblinear, and maximum iterations = 10,000.

Table 4. Hyperparameters values for the tuning process of LoGit.

Parameter	Value
Tolerance	$\{1 \times 10^{-4}, 1 \times 10^{-6}, 1 \times 10^{-8}\}$
C	$\{50,000, 100,000, 150,000\}$
Solver	$\{\text{lbfgs, saga, liblinear}\}$
Maximum iterations	$\{10,000, 50,000, 100,000\}$

Once the LoGit was trained, it was tested on 600 images it had not previously classified. The results show consistency in the classification within classes as well as the global classification accuracy, as shown in Table 5. The Matthews correlation coefficient is also acceptable at 0.864, indicating consistent results in all the confusion matrix categories, as shown in Figure 11. Figure 12 shows the corresponding ROC curve.

Table 5. Classification report of LoGit with subset of not previously seen images.

Label	Precision	Recall	F1-Score	Classification Accuracy	Count
good	0.912	0.952	0.932	0.932	294
bad	0.952	0.912	0.932		306

Matthews correlation coefficient: 0.864.

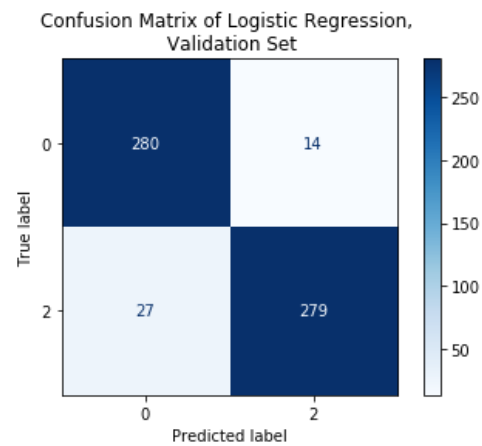


Figure 11. Confusion matrix of LoGit classification results on validation subset of 600 images.

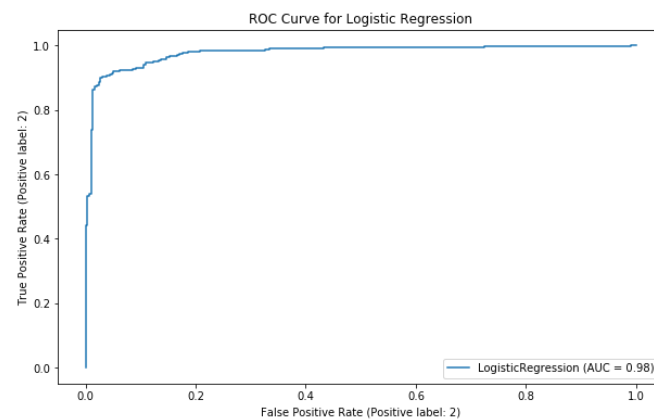


Figure 12. Receiver-operating characteristic curve for LoGit classification performance on validation subset of 600 images.

In order to bridge the numeric results to the clinical practice in the task of image quality assessment of eye fundus, a visualization of images adequately classified when

compared to ground truth labels is presented in Figure 13. Figure 14 presents examples of images where LoGit wrongly classified them according to the ground truth.

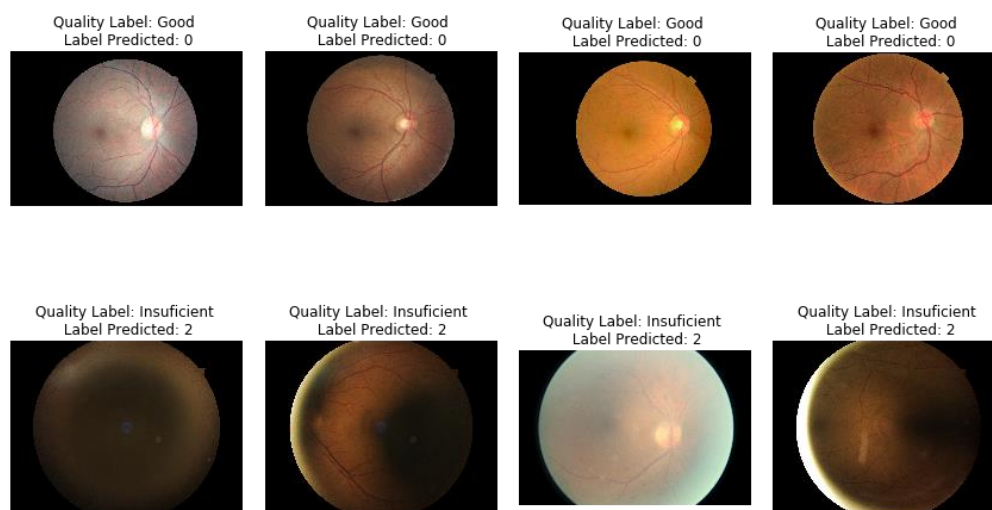


Figure 13. Receiver-operating characteristic curve for LoGit classification performance on validation subset of 600 images.

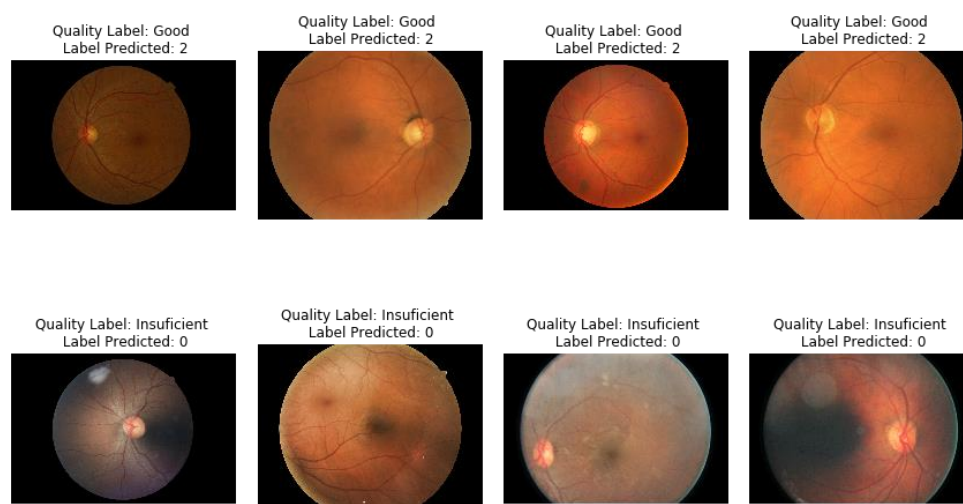


Figure 14. Subset of images wrongly classified by LoGit algorithm when compared to ground truth labels. For labels predicted value 0 = good, value 2 = bad.

On these visualizations, it is apparent that the topological descriptors are robust to changes in color, illumination, and blurring, despite being vectors conformed by 30 elements, all derived from vectorizations of representations of persistence diagrams.

5. Discussion

Digital funduscopy has become a valuable tool in the ophthalmic toolkit of contemporary clinical practice. With the advent of telemedicine and electronic medical records [47], funduscopy has moved to the mobile arena where now it is possible to perform it via smartphones [5,48]. With these relatively new approaches reaching the clinical practice, computer-aided diagnosis systems have also been proposed to assist the clinical process in situ or remotely [49,50].

One of the first steps in the digital fundus image analysis pipeline is the quality assessment of the image obtained, as presented in Section 1.2. This preprocessing step functions as a triage station allowing us to filter out those images not meeting the minimum quality needed to continue for clinical use [16,19,21,24]. This step is also performed by

the clinician, but it is trivial for humans, given how the brain processes and understands images. For machines, this step is fundamental and not trivial, since it will allow a smoother process down the pipeline of a CAD system.

As mentioned in Section 1.2, IQA techniques can be classified into three general groups, each representing the state of development in digital image analysis at the time of their publications (image characteristics, segmentation, and deep learning). To the knowledge of the authors, this is the first work using topological data analysis (TDA) to tackle the challenge of IQA in eye fundus images. The advantages of using TDA is that it represents less computational burden to the system, given that each image is first interpreted as a cubical complex and a cubical persistence calculation obtained represented by persistence diagrams, from which topological descriptors are extracted and vectorized. This allows us to run the classification task not on an image matrix but on a vector much smaller than the input data. Therefore, TDA, for the context of IQA in eye fundus images, contributes to an inherent dimensionality reduction in the data. This fact makes the method attractive, especially for contexts where limited computational power is of significance, such as when clinicians integrate mobile devices or telemedicine to their practice.

Another beneficial contribution is that the topological descriptors seem to be able to capture enough data from the image that a relatively simple and computationally not demanding algorithm suffices to render results with robust classification metrics; this has also been observed by [51]. Figure 13 shows evidence to the idea presented in the previous sentence, where the images in the examples present variability in their color, illumination, blurring, and anatomical landmarks. Still, the topological descriptors encapsulate sufficient information for the classifier to perform robustly. For the instances where the classification is not appropriate, it seems that the localization of the optic nerve, as well as a significant proportion of irregularity in color distribution in the same image, as well as blurring could explain the errors; nevertheless, more studies are needed in this direction to better understand this phenomenon.

Lastly, TDA seems to capture information on the form of topological descriptors, which allows for the classification to perform close to the ground truth labels, which in the context of images from clinical practice such as the EyePACS project, allows for the proposition of this method capturing information close to the way a clinician recollects image characteristics in order to decide if the image is of enough quality to be integrated in the clinical attention of the patient the data come from.

6. Conclusions

We propose a novel method for image quality assessment of eye fundus images based on the extraction of topological descriptors integrated into a machine learning classifier. The classification metrics are robust, and there is evidence that a topological approach facilitates the interpretation of the data in a similar way to how a clinician engages in these tasks during their practice. Further work is needed to investigate in more detail the clinical interpretations that might be suitable for the topological results obtained from eye fundus images. This could facilitate the use of TDA in later stages of the computer-assisted diagnosis pipeline in computational ophthalmological approaches.

Author Contributions: Conceptualization, G.J.A.-R. and J.I.N.-H.; formal analysis, G.J.A.-R.; funding acquisition, J.I.N.-H.; investigation, G.J.A.-R. and M.d.l.Á.C.-L.; methodology, G.J.A.-R., J.I.N.-H., M.d.l.Á.C.-L., and M.V.-B.; resources, J.I.N.-H.; software, G.J.A.-R.; supervision, J.I.N.-H., M.d.l.Á.C.-L., G.S.R.-C., J.d.D.S.-L., and P.R.-C.; visualization, G.J.A.-R.; writing—original draft, G.J.A.-R.; writing—review and editing, J.I.N.-H., M.d.l.Á.C.-L., G.S.R.-C., and J.d.D.S.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a doctoral scholarship by Consejo Nacional de Ciencia y Tecnología (CONACYT) in Mexico, the MyDCI Academic program, Facultad de Ingeniería Arquitectura y Diseño, and Universidad Autónoma de Baja California.

Institutional Review Board Statement: Ethical review and approval were waived for this study, given that the focus of the work is on analyzing digital images of human eye fundus previously gathered and published by another Institution.

Informed Consent Statement: Not applicable.

Data Availability Statement: For this study the publicly available eyePACS dataset was used. This data can be found here: <https://www.kaggle.com/c/diabetic-retinopathy-detection/data>.

Acknowledgments: The authors acknowledge the financial support from Universidad Autónoma de Baja California, the MyDCI graduate program, and the Mexican Consejo Nacional de Ciencia y Tecnología (CONACYT) for the realization of this project.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Burton, M.J.; Faal, H.B.; Ramke, J.; Ravilla, T.; Holland, P.; Wang, N.; West, S.K.; Bourne, R.R.A.; Congdon, N.G. Announcing the Lancet global health commission on global eye health. *Lancet Glob. Health* **2019**, *7*, e1612–e1613. [[CrossRef](#)]
- Bourne, R.R.A.; Seth, H.L.; Flaxman, R.; Braithwaite, T.; Cicinelli, M.V.; Das, A.; Jonas, J.B.; Keeffe, J.; Kempen, J.H.; Leasher, J. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis. *Lancet Glob. Health* **2017**, *5*, e888–e897. [[CrossRef](#)]
- Gordois, A.; Cutler, K.C.H.; Pezzullo, L.; Gordon, K.; Cruess, A.; Winyard, S.; Hamilton, W. An estimation of the worldwide economic and health burden of visual impairment. *Glob. Public Health* **2012**, *7*, 465–481. [[CrossRef](#)]
- WHO. *World Report on Vision*; WHO: Geneva, Switzerland, 2019.
- Rono, H.K.; Bastawrous, A.; Macleod, D.; Wanjala, E.; Di Tanna, G.L.; Weiss, H.A. Smartphone-based screening for visual impairment in Kenyan school children: A cluster randomised controlled trial. *Lancet Glob. Health* **2018**, *6*, e924–e932. [[CrossRef](#)]
- Mookiah, M.R.K.; Acharya, U.R.; Chua, C.K.; Lim, C.M.; Ng, E.Y.K.; Laude, A. Computer-aided diagnosis of diabetic retinopathy: A review. *Comput. Biol. Med.* **2013**, *43*, 2136–2155. [[CrossRef](#)]
- Yanoff, M. *Ophthalmic Diagnosis & Treatment*; JP Medical Ltd.: London, UK, 2014.
- Bruce, B.B. *Examining the Ocular Fundus and Interpreting What You See*; The American Academy of Neurology Institute: Minneapolis, MN, USA, 2017.
- Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **2016**, *316*, 2402–2410. [[CrossRef](#)]
- Voets, M.; Møllersen, K.; Bongo, L.A. Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *PLoS ONE* **2019**, *14*, e0217541. [[CrossRef](#)]
- Son, J.; Shin, J.Y.; Kim, H.D.; Jung, K.H.; Park, K.H.; Park, S.J. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology* **2020**, *127*, 85–94. [[CrossRef](#)] [[PubMed](#)]
- Krause, J.; Gulshan, V.; Rahimy, E.; Karth, P.; Widner, K.; Corrado, G.S.; Peng, L. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. *Ophthalmology* **2018**, *125*, 1264–1272. [[CrossRef](#)]
- Razzak, M.I.; Naz, S.; Zaib, A. Deep learning for medical image processing: Overview, challenges and the future. *Classif. BioApps* **2018**, 323–350.
- Abramoff, M.D.; Garvin, M.K.; Sonka, M. Garvin retinal imaging and image analysis. *IEEE Rev. Biomed. Eng.* **2010**, *3*, 169–208. [[CrossRef](#)]
- Bernardes, R.; Serranho, P.; Lobo, C. Digital Ocular Fundus Imaging: A Review. *Ophthalmologica* **2011**, *226*, 161–181. [[CrossRef](#)]
- Lalonde, M.; Gagnon, L.; Boucher, M.C. Automatic visual quality assessment in optical fundus images. *Vis. Interface VI2001* **2001**, *32*, 259–264.
- Davis, H.; Russell, S.; Barriga, E.; Abramoff, M.; Soliz, P. Vision-Based, Real-Time Retinal Image Quality Assessment. In Proceedings of the 2009 22nd IEEE International Symposium on Computer-Based Medical Systems, Albuquerque, NM, USA, 2–5 August 2009; pp. 1–6.
- Fleming, A.D.; Philip, S.; Goatman, K.A.; Olson, J.A.; Sharp, P.F. Automated assessment of diabetic retinal image quality based on clarity and field definition. *Investig. Ophthalmol. Amp Vis. Sci.* **2006**, *47*, 1120–1125. [[CrossRef](#)]
- Dias, J.M.P.; Oliveira, C.M.; da Silva Cruz, L.A. Retinal image quality assessment using generic image quality indicators. *Inf. Fusion* **2014**, *19*, 73–90. [[CrossRef](#)]
- Lee, S.C.; Wang, Y. Automatic Retinal Image Quality Assessment and Enhancement. *Med. Imaging 1999 Image Process.* **1999**, 3661, 1581–1590.
- Bartling, H.; Wanger, P.; Martin, L. Peter Wanger Automated Quality Evaluation of Digital Fundus Photographs. *Acta Ophthalmol.* **2009**, *87*, 643–647. [[CrossRef](#)]
- Niemeijer, M.; Abramoff, M.D.; van Ginneken, B. Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening. *Med. Image Anal.* **2006**, *10*, 888–898. [[CrossRef](#)]

23. Welikala, R.A.; Fraz, M.M.; Foster, P.J.; Whincup, P.H.; Rudnicka, A.R.; Owen, C.G.; Strachan, D.P.; Barman, S.A. Automated retinal image quality assessment on the uk biobank dataset for epidemiological studies. *Comput. Biol. Med.* **2016**, *71*, 67–76. [[CrossRef](#)] [[PubMed](#)]
24. Mahapatra, D.; Roy, P.K.; Sedai, S.; Garnavi, R. Retinal image quality classification using saliency maps and CNNs. In Proceedings of the International Workshop on Machine Learning in Medical Imaging, Athens, Greece, 17 October 2016; pp. 172–179.
25. Munkres, J.R. *Elements of Algebraic Topology*; CRC Press: Boca Raton, FL, USA, 2018.
26. Edelsbrunner, H.; Harer, J. *Computational Topology: An Introduction*; American Mathematical Soc.: Providence, RI, USA, 2010.
27. Chazal, F.; Michel, B. An introduction to topological data analysis: Fundamental and practical aspects for data scientists. *arXiv* **2017**, preprint. arXiv:1710.04019.
28. Carlsson, G. Topology and data. *Bull. Am. Math. Soc.* **2009**, *46*, 255–308. [[CrossRef](#)]
29. Cuadros, J.; Bresnick, G. EyePACS: An adaptable telemedicine system for diabetic retinopathy screening. *J. Diabetes Sci. Technol.* **2009**, *3*, 509–516. [[CrossRef](#)]
30. Fu, H.; Wang, B.; Shen, J.; Cui, S.; Xu, Y.; Liu, J.; Shao, L. Evaluation of retinal image quality assessment networks in different color-spaces. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; pp. 48–56.
31. Pérez, A.D.; Perdomo, O.; González, F.A. A Lightweight Deep Learning Model for Mobile Eye Fundus Image Quality Assessment. In Proceedings of the 15th International Symposium on Medical Information Processing and Analysis, Medellín, Colombia, 6–8 November 2020; Volume 11330.
32. Zomorodian, A. Topological data analysis. *Adv. Appl. Comput. Topol.* **2012**, *70*, 1–39.
33. Niethammer, M.; Stein, A.N.; Kalies, W.D.; Pilarczyk, P.; Mischaikow, K.; Tannenbaum, A. Analysis of blood vessel topology by cubical homology. In Proceedings of the International Conference on Image Processing, Rochester, NY, USA, 22–25 September 2002; Volume 2.
34. Pilarczyk, P.; Real, P. Computation of cubical homology, cohomology, and (co) homological operations via chain contraction. *Adv. Comput. Math.* **2015**, *41*, 253–275. [[CrossRef](#)]
35. Wagner, H.; Chen, C.; Vućini, E. Efficient computation of persistent homology for cubical data. In *Topological Methods in Data Analysis and Visualization II.*; Springer: Berlin, Germany, 2012; pp. 91–106.
36. Chung, M.K.; Bubenik, P.; Kim, P.T. Persistence diagrams of cortical surface data. In Proceedings of the International Conference on Information Processing in Medical Imaging, Williamsburg, VA, USA, 5–10 July 2009; pp. 386–397.
37. Atienza, N.; Escudero, L.M.; Jimenez, M.J.; Soriano-Trigueros, M. Persistent entropy: A scale-invariant topological statistic for analyzing cell arrangements. *arXiv* **2019**, arXiv:1902.06467.
38. Rucco, M.; Castiglione, F.; Merelli, E.; Pettini, M. Characterisation of the Idiopathic Immune Network Through Persistent Entropy. In Proceedings of the ECCS 2014; Battiston, S., De Pellegrini, F., Caldarelli, G., Merelli, E., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 117–128, ISBN 978-3-319-29226-7.
39. Efrat, A.; Itai, A.; Katz, M.J. Geometry helps in bottleneck matching and related problems. *Algorithmica* **2001**, *31*, 1–28. [[CrossRef](#)]
40. Kerber, M.; Morozov, D.; Nigmetov, A. Geometry helps to compare persistence diagrams. *arXiv* **2016**, arXiv:160603357.
41. Bubenik, P. Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.* **2015**, *16*, 77–102.
42. Bubenik, P.; Dłotko, P. A persistence landscapes toolbox for topological statistics. *J. Symb. Comput.* **2017**, *78*, 91–114. [[CrossRef](#)]
43. Reininghaus, J.; Huber, S.; Bauer, U.; Kwitt, R. A stable multi-scale kernel for topological machine learning. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 4741–4748.
44. Demšar, J.; Curk, T.; Erjavec, A.; Gorup, Č.; Hočevar, T.; Milutinovič, M.; Movžina, M.; Polajnar, M.; Toplak, M.; Staric, A.; et al. Orange: Data Mining Toolbox in Python. *J. Mach. Learn. Res.* **2013**, *14*, 2349–2353.
45. Jiao, Y.; Du, P. Performance measures in evaluating machine learning based bioinformatics predictors for classifications. *Quant. Biol.* **2016**, *4*, 320–330. [[CrossRef](#)]
46. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **2020**, *21*, 1–13. [[CrossRef](#)] [[PubMed](#)]
47. DeBry, P.W. Considerations for choosing an electronic medical record for an ophthalmology practice. *Arch. Ophthalmol.* **2001**, *119*, 590–596. [[CrossRef](#)] [[PubMed](#)]
48. Barikian, A.; Haddock, L.J. Smartphone assisted fundus funduscopy/photography. *Curr. Ophthalmol. Rep.* **2018**, *6*, 46–52. [[CrossRef](#)]
49. Díez-Sotelo, M.; Díaz, M.; Abalde, M.; Gómez-Ulla, F.; Penedo, M.G.; Ortega, M. A novel automatic method to estimate visual acuity and analyze the retinal vasculature in retinal vein occlusion using swept source optical coherence tomography angiography. *J. Clin. Med.* **2019**, *8*, 1515. [[CrossRef](#)] [[PubMed](#)]
50. Arsalan, M.; Owais, M.; Mahmood, T.; Cho, S.W.; Park, K.R. Aiding the diagnosis of diabetic and hypertensive retinopathy using artificial intelligence-based semantic segmentation. *J. Clin. Med.* **2019**, *8*, 1446. [[CrossRef](#)]
51. López-Reyes, V.; Cosío-León, M.A.; Avilés-Rodríguez, G.J.; Martínez-Vargas, A.; Romo-Cárdenas, G. A topological approach for the pattern analysis on chest X-Ray images of COVID-19 patients. In *Medical Imaging 2021: Physics of Medical Imaging*; International Society for Optics and Photonics: Bellingham, WA, USA, 2021; Volume 11595.