

Impact of Keypoint Detection on Graph-based Characterization of Blood Vessels in Colonoscopy Videos

Joan M. Núñez, Jorge Bernal, Miquel Ferrer, and Fernando Vilariño

Computer Vision Center (Universitat Autònoma de Barcelona), Edificio O, Campus UAB, 089193 Bellaterra (Cerdanyola), Barcelona, Spain

jmnunez@cvc.uab.es
<http://www.cvc.uab.es/>

Abstract. We explore the potential of the use of blood vessels as anatomical landmarks for developing image registration methods in colonoscopy images. An unequivocal representation of blood vessels could be used to guide follow-up methods to track lesions over different interventions. We propose a graph-based representation to characterize network structures, such as blood vessels, based on the use of intersections and endpoints. We present a study consisting of the assessment of the minimal performance a keypoint detector should achieve so that the structure can still be recognized. Experimental results prove that, even by achieving a loss of 35% of the keypoints, the descriptive power of the associated graphs to the vessel pattern is still high enough to recognize blood vessels.

Keywords: Colonoscopy, Graph Matching, Biometrics, Vessel, Intersection

1 Introduction

Colorectal cancer (CRC) is nowadays the fourth cause of cancer death worldwide and its survival rate depends on the stage CRC is detected. Early detection by means of efficient colon screening is crucial to reduce CRC mortality. Colonoscopy is considered the gold standard for colon screening although it presents some drawbacks, such as the fact that some polyps are still missed the most relevant of them [1].

Intelligent systems have been created to provide additional information either in intervention time or in post-intervention. The most straightforward application is the development of computer-aided diagnosis methods (CAD) although other applications, like the assessment of the quality of a single colonoscopy intervention. One of these potential applications is the development of patient's follow-up methods to allow the recognition of a single area of the colon containing a lesion when that area is revisited. Consequently we need to define and characterize those anatomical structures that remain stable over time in order to use them as markers to guide these follow-up methods. The only elements of the

endoluminal scene whose appearance tend to keep stable are blood vessels and polyps -if they are not removed during biopsy-, as the appearance of structures such as folds is more prone to change.

In this paper we focus on the characterization of blood vessels as anatomical landmarks with potential to be used as part of the development of follow-up methods. Blood vessels can be seen as branching patterns. The characterization of branching patterns has been deeply studied in other research fields, as the presence of these kind of structures is very common either in nature -palm prints [2]- or in anatomical structures -vascular trees [3]-. An accurate detection of these patterns along with a proper characterization of the network properties plays a key role for applications using this kind of information. The segmentation of vessel patterns can be a difficult task though, given the nature of the procedure or image quality issues, such as resolution. However, keypoints in the pattern can be used to unequivocally characterize branching structures without the necessity of computing an accurate segmentation of the vessel pattern. These keypoints can potentially be identified as anatomical landmarks to be used in image registration methods.

Therefore, an accurate detection of these keypoints appears as a key stage for a good performance of these methods. Basically, there are two kinds of keypoints junctions/intersections and endpoints. Detection of vascular intersections has been mainly studied in the field of retinal imaging. Available methods have been separated in two categories: *geometrical feature-based* and *model-based* methods [4, 3].

The first category groups methods which commonly start by a pixel-level processing stage followed by post-processing analysis specific for each methods. Methods belonging to this group tend to involve adaptive filtering and branch analysis based on thinned structures -being thinning a common step in the majority of available methods and an important source of error-. Some examples of *geometrical feature-based* methods appear in the works of Bhuiyan et al. [4] or Saha et al. [5]. Regarding the second category, *model-based* methods are based on a vectorial tracing of the desired structure. These methods use seed points as initial locations so vascular structures that appear in the image can be recursively tracked from them. Examples of this can be found in [6, 7]. Finally, the method of [3] offers a completely different approach which is based on the definition of COSFIRE filters to detect intersections in retinal images.

At this point, one question arises: which is the minimal performance these methods have to achieve so that those keypoints can be used as reliable anatomical landmarks. That is, if we characterize blood vessels in terms of intersections and/or endpoints, how many of them could we miss and still be able to recognize a posterior appearance of the same structure in a different frame. In this paper we propose a validation framework aiming to assess whether the performance of a given intersection detection method is enough to characterize blood vessels.

We represent blood vessels using graphs. Graphs consist of a finite set of nodes connected by edges and they are one of the most general data structure in computer science. Due to the ability of graphs to represent properties of entities

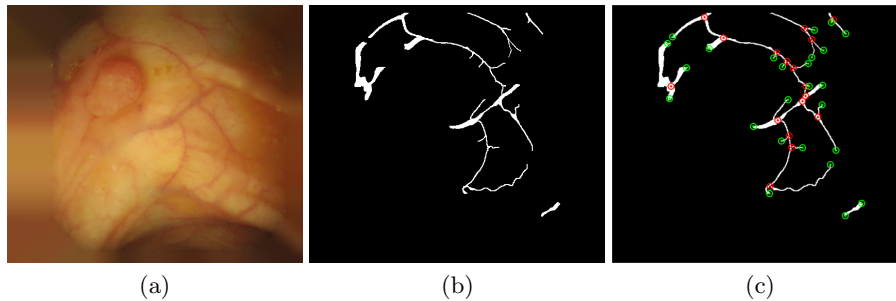


Fig. 1: Keypoint definition in colonoscopy images. (a) Original image. (b) Binary representation of blood vessels. (c) Keypoint characterization: intersection/junctions (red), endpoints (green).

and binary relations at the same time, a growing interest in graph-based object representation can be observed in various fields. In bio and chemoinformatics, for instance, graph based representations are intensively used [8]. Further areas of research where graph based representations draw attention are web content and data mining [9], image classification [10], and graphical symbol and character recognition [11], among others.

We use a graph representation in which nodes can be either intersections/junctions or endpoints. Our validation will assess the similarity -using graph edit distance- between the original graph and an altered version created by removing some nodes from the original graph (simulating that those keypoints are not detected). We do not intend to propose the best graph matching framework but to assess the descriptive power of vessel content graphs when a percentage of nodes are removed. The validation is performed on a database of 40 colonoscopy images specially rich in blood vessel content.

After this introduction, we present in Section 3 our strategy to create and compare graphs from detected intersections. The experimental setup is introduced in Section 4. Experimental results are exposed in Section 5. We close this paper with the conclusions and future work in Section 6.

2 Vascular structures in colonoscopy images

Blood vessels appear as vascular structures in colonoscopy images. As mentioned before, they can be seen as branching patterns and different keypoints can be used to help in their unequivocal characterization. In our case we define two different keypoints associated to blood vessels: intersections/junctions and endpoints. The definition of this keypoints can be better understood by looking at a binary segmentation of blood vessels. Considering this, a single pixel is labelled as an intersection/junction if it was identified as the point where at least three of the branches meet together. Conversely a pixel is labelled as an endpoint if it constitutes the end of a given branch and not reaching the boundary of informative area of the colonoscopy image.

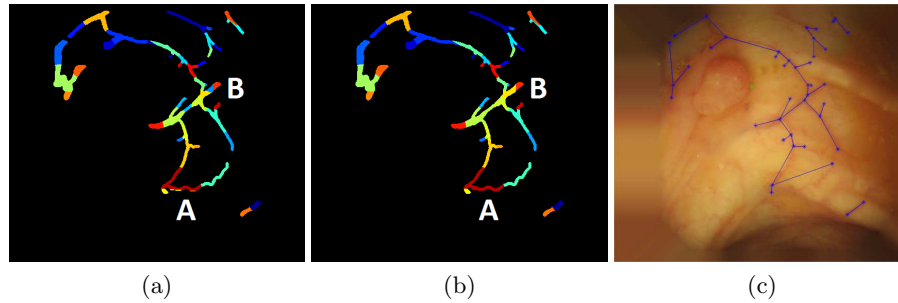


Fig. 2: Example of adjacency matrix calculation to create the final graph. (a) First segmentation. (b) Refined segmentation. (c) Final graph. Areas marked with A and B exemplify problems related with the first graph segmentation.

Figure 1 shows an example of manually labelled keypoints in a vessel pattern. Once the keypoints are defined and characterized for each colonoscopy image with blood vessels we can proceed with the extraction of a graph.

3 Graph matching strategy

The kind of blood vessels we are dealing with contain high structural information. Junctions and endpoints and the way they relate to other points appear to be crucial for blood vessel characterization. For this reason, we use a graph matching framework to assess the impact of the selection of nodes in the robustness of a graph as a characterization of a vascular pattern. To accomplish this task, we first need to transform our images into graphs and define the attributes of both the nodes and the edges. Once the graphs are constructed, a similarity measure to compare such graphs is needed. In our case we use the graph edit distance. We remark that the purpose of this study is the assessment of the consequences of losing keypoint information in the descriptive power of the resulting vessel pattern graphs.

3.1 Graph extraction

We extract the graph given a set of keypoints and the binary pattern to be characterized. The keypoints will stand for the nodes and the binary pattern provides the structural information to create the adjacency matrix of the graph. The computation of the adjacency matrix from the binary pattern comprises the following steps (see example in Figure 2):

- In order to find adjacencies between branches, a first segmentation of the graph is achieved by grouping all the pixels behind the binary pattern into clusters represented by each keypoint detected. The criteria to assign a given pixel to a cluster is the Euclidean distance -Figure 2a-.

- We can observe from Figure 2a how the first segmentation may present some incoherences as some pixels and branches might be associated to keypoints which do not have connectivity. To solve this problem, all the regions which do not contain a keypoint are merged into any of the regions that are connected to them -Figure 2b-.
- The final graph is the region adjacency graph extracted from the segmented image -Figure 2c-.

The distance of each node to the centroid of the vessel pattern is associated as the only attribute of each graph node. The distances are normalized with respect to the longer distance. This makes this simple characterization invariant to rotation and scale. This decisions intend to lead to an experiment framework to validate the keypoint impact independently from other consideration.

3.2 Graph edit distance

Graph edit distance [12, 13] is one of the most flexible and versatile approaches to error-tolerant graph matching. One of the major advantages of graph edit distance is that it is able to cope with directed and undirected graphs, as well as with labelled and unlabelled graphs. If there are labels on nodes, edges, or both, no constraints on the respective label alphabets have to be considered. Moreover, through the concept of cost functions, graph edit distance can be adopted and tailored to various applications such as fingerprint classification [14], diatom identification [15], or clustering of color images [16], just to mention a few.

The major drawback of graph edit distance is its high computational complexity that restricts its applicability to graphs of rather small size. In fact, graph edit distance belongs to the family of quadratic assignment problems (QAPs), which in turn belong to the class of NP-complete problems. That is, an exact and efficient algorithm for the graph edit distance problem can not be developed unless $P = NP$. Therefore, both the development of fast approximation algorithms and the gradual improvement of established approximation schemes are important and reasonable lines of research. In recent years, a number of methods addressing the high computational complexity of graph edit distance computation have been proposed [17, 18, 19, 20].

We propose the use of the algorithmic framework presented in [21] which allows the approximate computation of graph edit distance in a substantially faster way than traditional methods. The basic idea of this approach is to reduce the difficult problem of graph edit distance to a linear sum assignment problem (LSAP). LSAPs basically constitute the problem of finding an optimal assignment between two independent sets of entities. There is a large amount of available polynomial algorithms for LSAPs and an interested reader can find more information in [22].

Given two graphs, the source graph g_1 and the target graph g_2 , the basic idea of graph edit distance is to transform g_1 into g_2 using some edit operations. A standard set of distortion operations is given by insertions, deletions, and

substitutions of both nodes and edges. A sequence of edit operations e_1, e_2, \dots, e_k that transform g_1 completely into g_2 is called an edit path between g_1 and g_2 . To find the most suitable edit path out of all possible edit paths between two graphs g_1 and g_2 , a cost is introduced for each edit operation, measuring the strength of the corresponding operation. The edit distance of two graphs is then defined by the minimum cost edit path between two graphs. As can be thought, the cost function is highly dependent on the attributes of the nodes and edges. A different cost, specific for each problem to be solved, is applied to each of the distortion operations.

4 Experimental setup

We have created a database named COLON-VESSEL where to assess the minimum performance that a given keypoint detection method should achieve to ensure reliable blood vessels characterization. COLON-VESSEL database has been created from 15 different colonoscopy videos belonging to *CVC_COLON_DB* [23]. These videos were obtained from St. Vincent’s Hospital and Beaumont Hospital in Dublin, Ireland.

An expert selected 40 frames of size 574×500 specially rich in terms of vascular information. A double ground truth consisting of a mask of the blood vessels along with the position of the intersections/junctions and endpoints was provided for each of the 40 frames. With respect to keypoint annotation, a single pixel was labelled as an junction if it was identified as the point where at least three of the branches meet together. Conversely a pixel was labelled as an endpoint if it constitutes the end of a given branch and not reaching the boundary of the image. The number of junctions and endpoints per image ranges from 9 to 150 and 9 to 150, respectively. We can observe an example of the content of the database along with the ground truth in Figure 1.

We run several experiments to assess the degree of robustness of blood vessel representation using graphs. The graphs are created from the ground truth provided by experts. From the original image graph, we progressively and randomly eliminate intersections/junctions or endpoints -and the corresponding edges converging to them-. A query consisting of the identification of the altered graph over the set of the original graphs for each image is performed. The experiment removes a certain percentage of keypoints from the original graph (Figure 3 shows an example) before they are compared in terms of graph edit distance. This proposed graph matching framework will provide the closest graph among the data set for each query frame. Hence the system will always provide a matching -correct or not-. We do not address this issue as our goal is to assess the descriptive power of keypoint graphs rather than proposing a real solution to the frame matching problem, which should be addressed in the future.

For the particular case of blood vessel structure representation using graphs we define the following cost function:

- Node deletion/insertion: cost is a constant value equal to 0.9.

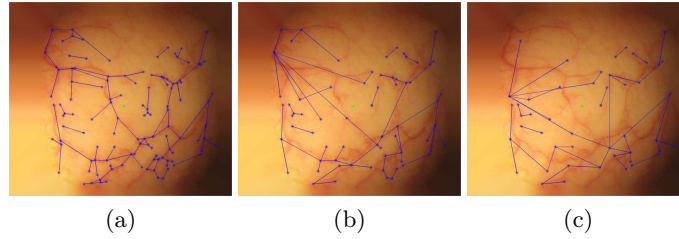


Fig. 3: Graph modification examples (note that the removed keypoints are different as they are removed randomly). (a) Original graph. (b) Graph after 40 nodes removal. (c) Graph after 70 nodes removal.

- Node substitution: absolute value of the difference between distances to the centroid.
- Edge deletion/insertion: cost is a constant value equal to 1.7.
- Edge substitution: cost is 0.

TODOTODOTODOTODO: Brief on consequences of the cost function regarding the paper goal and contribution. what if higher or lower, etc

The removal of keypoints is carried out in steps of a 5% -which entails 20 intervals- and each step is repeated 10 times to provide statistically significant results regarding the impact of the removal of keypoints. We run three different experiments, regarding the keypoints we used -junctions, endpoints and both junctions and endpoints- to assess the descriptive power of each possibility (see examples in Figure 4).

5 Results

Figure 5 shows two results for the three experiments introduced above.

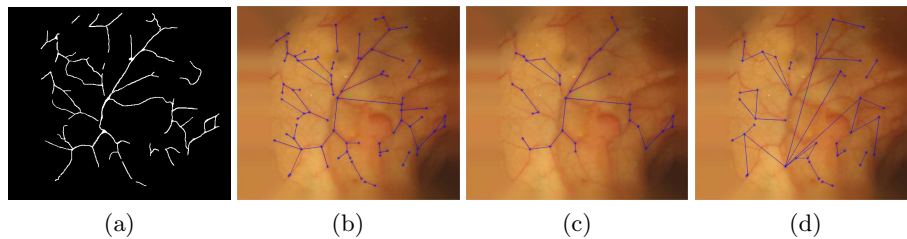


Fig. 4: Graph extraction examples. (a) Binary pattern. (b) Graph created with junctions and endpoints on original image. (c) Graph created with junctions on original image. (d) Graph created with endpoints on original image.

Regarding intersections/junctions, we observe in Figure 5a that a 100% matching between incomplete graphs is possible if a given intersection detector is able to achieve less than 35% missing error. Matching success decreases when we remove more than 35% of the intersections in the image, reaching very low matching performance when removal reaches 65%. As mentioned above, we do not obtain a 0% of matching because the systems always provides the closes graphs. Hence when all the nodes are removed the similarity output will provide with a match with the graph with less nodes. Figure 5b presents a breakdown of the results for each image; this experiment confirms that for the majority of the images matching starts to fail when we remove more than 35% of the intersections. We can observe that in the case of images with less nodes matching fails when a slightly higher percentage of intersections is removed.

Regarding endpoint removal, we can observe a similar behaviour in Figure 5c although in this case the minimal error permitted for keypoint detectors is moderately higher -around 40%-. One possible reason behind this differences is that in this case nodes tend to have less edges reaching them. Results broken down per image -Figure 5d- show again the same trend, although in this case there seems to be a more direct relation between the number of endpoints in the image and the threshold percentage needed to lose matching precision.

Finally, we present in Figure 5e a plot showing the behaviour when we remove randomly either a junction or a endpoint. In this case the threshold percentage is close to the one achieved by junctions, which can be associated to having more junctions than endpoints in a given image.

Figure 6 shows -for the three experiments- the distance between pairs of equivalent graphs in four cases with a different amount of keypoints removed -we show results with 100% of node removal to fix the maximum possible distance that we can have between the original and the altered graph-. We can see in Figure 6a that there is a direct correspondence between the number of nodes in the image and the graph edit distance, decreasing the latter for images where the number of nodes is minimal. This trend is kept for the cases of endpoints -Figure 6b- and the indistinct elimination of junctions or endpoints -Figure 6c-. We can conclude from these results that the more nodes we remove, the higher is the cost to transform the original graph into the altered graph. High alterations of the original graph -resulting in high graph edit distances- are proven to have an impact in matching success, as observed in Figure 5. Moreover, we can observe that there is a higher impact associated to endpoints on the graph edit distances between image pairs, which confirms the descriptive power of endpoints in graphs representing blood vessels structures.

6 Conclusions

One of the needs expressed by physicians is the ability of following-up lesions in colonoscopy procedures. One of the elements present in the endoluminal scene that can be used to help tracking these lesions are blood vessels, as their appearance tends to keep stable along different revisions of the same patient. Blood

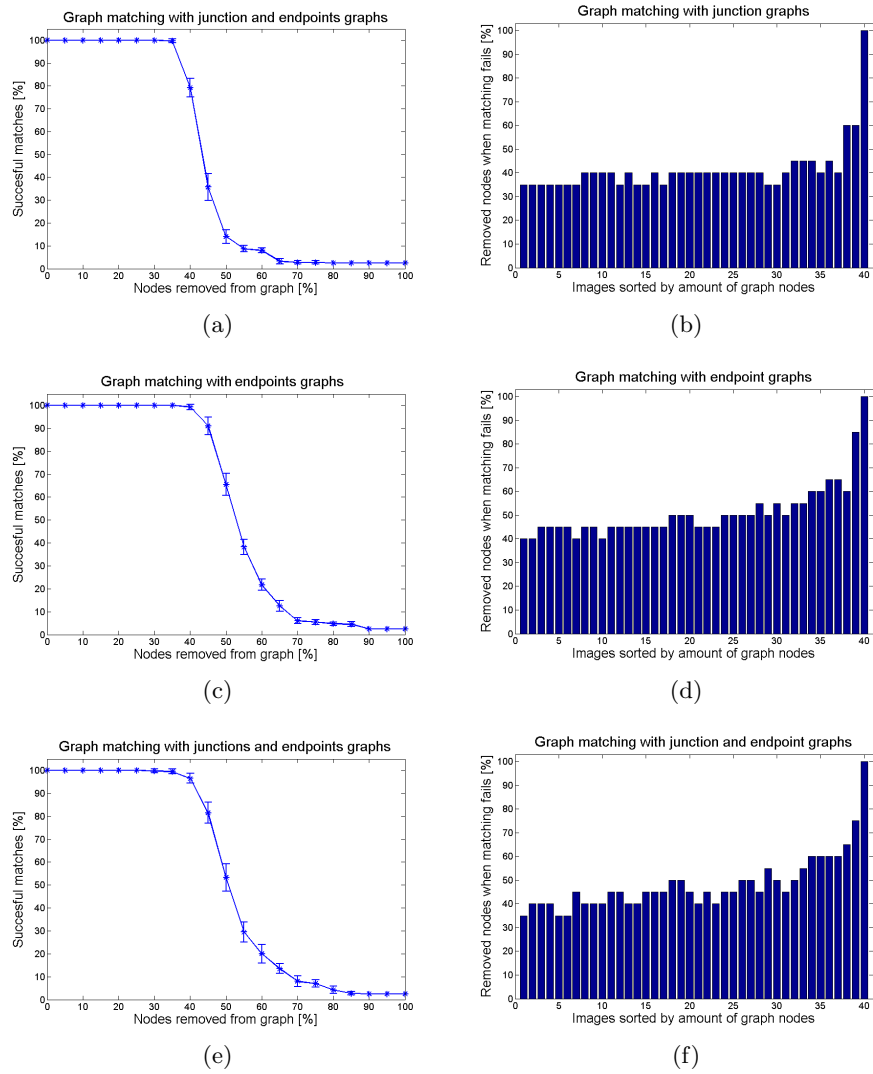


Fig. 5: Impact of percentage of node removal in the number of correct matches. (a,b) Junctions. (c,d) Endpoints. (e,f) Junctions and endpoints.

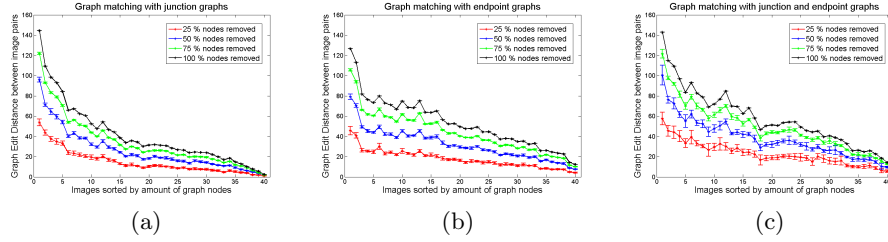


Fig. 6: Study of graph edit distance according to the percentage of nodes eliminated: (a) Junctions. (b) Endpoints. (c) Junctions and endpoints.

vessels can be seen as branching patterns and therefore they may be characterized by means of the position and number of branching points. Our approach for characterizing blood vessels consists of using graphs created from detected branching points in a way such their structure can be unequivocally recognized.

In this paper we present a study of the impact of the accuracy in keypoint detection in the way to develop graph matching based registration systems for colonoscopy frames. We propose a graph matching configuration whose only purpose is to provide a framework for our experiments. Experimental results show that there is a clear decrease in the success of the matching framework when a considerable amount of nodes is removed. We confirm that, when the removal of nodes keeps below certain percentages, the descriptive power of the resulting graphs is high enough to provide with successful matches. This conclusion applies -with small differences- for the three kinds of graph characterization we have tested: using intersections/junctions, endpoints, or both intersections and endpoints.

Future work should consist of making further analysis regarding the impact of the connectivity of a given keypoint, that is, the relationship between missing the detection of a highly connected keypoints and the descriptive power of the associated graph. Given the trends concluded from this study, further research should be performed regarding bigger datasets and considering the several kind of image deformation that pattern can suffer in a colonoscopy procedure. As keypoints are the source of the graphs characterizing blood vessels, an accurate intersection and endpoint detector will be necessary. The degree of information in a query frame should also be evaluated to test the ability of a matching system to provide a matching for the query or to discard the request because of lack of information in the scene.

Another interesting research line would address the assessment of the descriptive power of graphs after a partial removal of node information in other kind of graphs from structured patterns.

7 Acknowledgements

This work was supported in part by the Spanish Gov. grants TIN2012-33116, MICINN TIN2009-10435 and the UAB grant 471-01-2/2010.

References

- [1] B. Bressler, L. Paszat, Z. Chen, D. Rothwell, C. Vinden, and L. Rabeneck, “Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis,” *Gastroenterology*, vol. 132, no. 1, pp. 96–102, 2007.
- [2] M. Pudzs, R. Fuksis, and M. Greitans, “Palmprint image processing with non-halo complex matched filters for forensic data analysis,” in *Biometrics and Forensics (IWBF), 2013 International Workshop on*, pp. 1–4, IEEE, 2013.
- [3] G. Azzopardi and N. Petkov, “Automatic detection of vascular bifurcations in segmented retinal images using trainable cosfire filters,” *Pattern Recognition Letters*, vol. 34, no. 8, pp. 922–933, 2013.
- [4] A. Bhuiyan, B. Nath, J. Chua, and K. Ramamohanarao, “Automatic detection of vascular bifurcations and crossovers from color retinal fundus images,” in *Signal-Image Technologies and Internet-Based System, 2007. SITIS’07. Third International IEEE Conference on*, pp. 711–718, IEEE, 2007.
- [5] S. Saha and N. Dutta Roy, “Automatic detection of bifurcation points in retinal fundus images,” *Latest Research in Science and Technology. International Journal of*, vol. 2, no. 2, pp. 105–108, 2013.
- [6] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum, “A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 3, pp. 347–364, 2002.
- [7] C.-L. Tsai, C. V. Stewart, H. L. Tanenbaum, and B. Roysam, “Model-based method for improving the accuracy and repeatability of estimating vascular bifurcations and crossovers from retinal fundus images,” *Information Technology in Biomedicine, IEEE Transactions on*, vol. 8, no. 2, pp. 122–130, 2004.
- [8] P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret, and J.-P. Vert, “Graph kernels for molecular structure-activity relationship analysis with support vector machines,” *Journal of chemical information and modeling*, vol. 45, no. 4, pp. 939–951, 2005.
- [9] A. Schenker, H. Bunke, M. Last, and A. Kandel, “Graph-theoretic techniques for web content mining,” 2005.
- [10] Z. Harchaoui and F. Bach, “Image classification with segmentation graph kernels,” in *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*, pp. 1–8, IEEE, 2007.
- [11] J. Lladós and G. Sanchez, “Graph matching versus graph parsing in graphics recognition: a combined approach,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, no. 03, pp. 455–473, 2004.
- [12] A. Sanfeliu and K.-S. Fu, “A distance measure between attributed relational graphs for pattern recognition,” *Systems, Man and Cybernetics, IEEE Transactions on*, no. 3, pp. 353–362, 1983.
- [13] H. Bunke and G. Allermann, “Inexact graph matching for structural pattern recognition,” *Pattern Recognition Letters*, vol. 1, no. 4, pp. 245–253, 1983.
- [14] M. Neuhaus and H. Bunke, “A graph matching based approach to fingerprint classification using directional variance,” in *Audio-and Video-Based Biometric Person Authentication*, pp. 191–200, Springer, 2005.

- [15] R. Ambauen, S. Fischer, and H. Bunke, “Graph edit distance with node splitting and merging, and its application to diatom identification,” in *Graph Based Representations in Pattern Recognition*, pp. 95–106, Springer, 2003.
- [16] A. Robles-Kelly and E. R. Hancock, “Graph edit distance from spectral seriation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 3, pp. 365–378, 2005.
- [17] M. C. Boeres, C. C. Ribeiro, and I. Bloch, “A randomized heuristic for scene recognition by graph matching,” in *Experimental and Efficient Algorithms*, pp. 100–113, Springer, 2004.
- [18] S. Sorlin and C. Solnon, “Reactive tabu search for measuring graph similarity,” in *Graph-Based Representations in Pattern Recognition*, pp. 172–182, Springer, 2005.
- [19] D. Justice and A. Hero, “A binary linear programming formulation of the graph edit distance,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 8, pp. 1200–1214, 2006.
- [20] M. Neuhaus, K. Riesen, and H. Bunke, “Fast suboptimal algorithms for the computation of graph edit distance,” in *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 163–172, Springer, 2006.
- [21] K. Riesen and H. Bunke, “Approximate graph edit distance computation by means of bipartite graph matching,” *Image and Vision Computing*, vol. 27, no. 7, pp. 950–959, 2009.
- [22] R. Burkard, M. Dell’Amico, and S. Martello, *Assignment Problems, Revised Reprint*, vol. 106. SIAM, 2012.
- [23] J. Bernal, J. Sánchez, and F. Vilariño, “Towards automatic polyp detection with a polyp appearance model,” *Pattern Recognition*, vol. 45, no. 9, pp. 3166–3182, 2012.