Analysis and Implementation of Methodologies for the Monitoring of Changes in Eye Fundus Images

A. Gelroth, D. Rodríguez, A. Salvatelli, B. Drozdowicz, G. Bizai

Artificial Intelligence Group, Facultad de Ingeniería, Universidad Nacional de Entre Ríos (UNER).

E-mail: asalvatelli@bioingenieria.edu.ar,

Abstract. We present a support system for changes detection in fundus images of the same patient taken at different time intervals. This process is useful for monitoring pathologies lasting for long periods of time, as are usually the ophthalmologic. We propose a flow of preprocessing, processing and postprocessing applied to a set of images selected from a public database, presenting pathological advances. A test interface was developed designed to select the images to be compared in order to apply the different methods developed and to display the results. We measure the system performance in terms of sensitivity, specificity and computation times. We have obtained good results, higher than 84% for the first two parameters and processing times lower than 3 seconds for 512x512 pixel images. For the specific case of detection of changes associated with bleeding, the system responds with sensitivity and specificity over 98%.

1. Introduction

Today, the application of new medical technologies is a reality that does not cover all practices and does not extend to all professionals. Moreover, although many of these professionals have access to computer systems with large volumes of information, these systems can be underutilized if they are not assisted with tools, to deal in a practical, effective, efficient and synergic way for their business.

In the case of diagnosis and treatment of eye diseases, it is a routine practice to take photographs of the fundus, which are then analyzed in order to find characteristic signs and symptoms, thus evaluating the possible pathological development. In the analysis of the evolution of ocular pathology, by comparing images of the same patient taken at different time intervals, it is unusual for specialists to have computerized assistance to facilitate the observation of these images simultaneously and even less to assist in distinguishing changes.

We study methods that take advantage of digital images acquired at different time intervals, to change detection. Therefore we aim to assist professional ophthalmologists, through the computer system, by facilitating the monitoring of their patients and achieving a positive impact on treatment.

The selected methodologies are part of a test system that will allow us to analyze the results of its implementation on a series of images.
2. Materials and Methods

2.1. Materials
Based on online public databases (Neuro-Ophthalmology Collection Dr. William F. Hoyt, Atlas of eye disorders "RedAtlas.org - Recognizing Eye Disease: A Visual Review of Ophthalmic Disorders") pairs of images (each associated with a patient) were selected, taking into account clinically significant changes. These images were marked by specialists with medical practice, disregarding signs and diseases of low incidence. From this process, about 40 pairs of images were selected. Images generally come in two formats, JPEG and TIFF, RGB color with 8 bits for each of the 3 channels. The image sizes and resolutions range from 100 KBytes to 11 MBytes, with sizes of 128 x 128 to 2200 x 2200 pixels respectively.

We developed a graphical user interface that was necessary for the purposes of selection, combination and configuration of the methodologies to analyze.

2.1.1. Test images description
The set of images previously employed in the demarcation of the Gold Standart Masks (GSM) [1, 2] for parameter setting was grouped with a new set of pairs of images to form the "Test Group", comprising 9 pairs of images in total and more than 100 regions subject to change.

In the pairs of images belonging to the test group, changes associated with bleeding were predominant, following those associated with changes in the vessels, and other showed exudates. Most of these changes were found mainly in the area corresponding to the retina itself because of their greater size compared to the optic disc. Images with relative displacement greater than 20% in the photographic capture were discarded. At the same time changes in images containing edges of the optic disc were not considered, suggesting for these cases other methodologies with a different approach. Pairs of images associated with pathological signs that radically altered the nature of the retina (as can be the neoplasia and associated malignancies) were not considered.

2.2. Methods
The main considerations related to the selection of methodologies are directed toward a future development of an assistive system which can support ophthalmologists in the task of change detection in series of images.

2.2.1. Registration: Application of transformation functions with manual selection of point.
A set of pairs of points were generated manually, taking physical correspondence between the images that were part of each pair of fundus images [3]. One of the main advantages of this methodology lies in the simplicity of the method and the absence of problems, typical of automatic methods associated to identifying features when these are very distorted because of the pathological nature of the image. From the pairs of points generated various transformation functions were applied [4-7], obtaining very good results with the affine and polynomial transformations. It was observed that the affine transformation required fewer pairs of points that the polynomial and was approximately 50% faster.

2.2.2. Masking relative to the size of the image
Masks were generated as circles centered on the center of the image, with diameters predefined by the user as a percentage of the dimensions of the image. This methodology allows us to separate the masking processes from the quality of the image and possible pre-masking.
2.2.3. **Illumination correction - homomorphic filtering**

The applications of this methodology [8-11] on the set of images gave good results. This methodology largely reduced the effects of nonuniform illumination distribution.

These good results were obtained for high-resolution images that were presented in Table I. However, this technique showed some problems associated with the presence of artifacts uniformly distributed in the image with high intensity values.

| Stage               | 256x256 | 512x512 | 1345x1134 | 2048x2048 |
|---------------------|---------|---------|-----------|-----------|
| Homomorphic Filtering | 0.75    | 0.98    | 11.7      | 53.4      |

2.2.4. **Illumination correction – Ratio of green channel and red channel (G / R)**

The implementation of this methodology requires the preparation of images, to prevent zero division, and aims to override the lighting components that were present in green and red channels, generating a grayscale image in which pixels intensities were weighted that more differences exhibited between these channels [2.12 to 15].

Good results were obtained when the lighting components were minimized, with excellent processing times for this stage. These are presented in Table II.

| Stage               | 256x256 | 512x512 | 1345x1134 | 2048x2048 |
|---------------------|---------|---------|-----------|-----------|
| Ratio G / R         | 0.01    | 0.06    | 0.3       | 0.9       |

2.2.5. **Global statistical normalization**

This technique, based on the definition of identical mean values and standard deviation for all images used, produced good results for the following stages.

Here, it was observed that the use of this technique allowed the use of parameters such as threshold value for binarization stage, which were practically independent from the pair of images used. Thus, these parameters could be defined solely by the combination of methodologies used in different stages of processing [16-18].

2.3. **Processing**

2.3.1. **Calculation of the Image Difference - Simple Difference**

This technique consists of subtracting from a reference image a second image of the same scene, in which there may be changes. Algorithms were implemented to obtain the absolute value of subtraction and solve potential format problems.

2.3.2. **Binarization Image Difference - Fixed threshold**

The use of a fixed threshold is valid only when it is preceded by an appropriate standardization process. The process of defining the threshold was based, for each combination of methodologies, from masks generated by the expert, finding values that vary slightly for all test pairs.

2.4. **Post Processing**

2.4.1. **Processing mask change - regions detection and filter properties.**

We implemented the detection and segmentation of clusters of pixels in the changes mask [19]. Then features such as its centroid, number of pixels, bounding boxes, etc, were extracted.
Based on the number of pixels associated with each area it is possible to remove the masking areas that do not exceed a minimum. This threshold was defined based by experience and with the help of the MGS generated by the expert. In this way it was possible to remove isolated pixels associated with noise, without removing or modifying areas associated with significant changes, capacity that other methodologies, such as binary operators do not have [12.19-20]. Thresholds are defined in relation to the size of the images because the regions of change will grow proportionately to the size of the images to be processed and, therefore, the thresholds will adapt to those changes (for details see the final flow diagram).

2.5. Results presentation.
We have chosen to overlap the original images with the perimeter of areas with detected changes, so that we can compare these to the areas drawn from the MGS generated by the expert. Also, changes could be demarcated as areas filled through the use of bounding boxes [21-22]. The latter help to define areas of changes rather than changes, which would facilitate the specialist to evaluate in the same image the significant signs without obstruct her/his observation.

2.6. Workflow final
In Figure 1, we present the process of transforming the input images along the workflow, resulting in final change mask. This workflow embedded the image with the numbered regions from the mask generated by the system, which was provided to the expert.

2.7. Performance measurement methodology
Each region detected by the system was enumerated manually. Furthermore, for each pair of images, a template was generated. This, together with pairs of images and their corresponding masks demarcated by the system, was referred to the specialist to be completed, as indicated in Table III and IV.

| Image N. | Region | Sign | Type of change | Subtype if there | Properly Change Detected | Improperly Change Detected | Change Detected |
|----------|--------|------|----------------|------------------|---------------------------|---------------------------|-----------------|
| 1        | H      | A    | Perif.         | x                |                           |                           |                 |
| 2        | E      | F    |                |                 |                           |                           | x               |
| 3        | Otros  | A    |                |                 |                           |                           | x               |

| Non Detect Change | H | E | V | O | R | Otros |
|-------------------|---|---|---|---|---|--------|
Figure 1. Workflow final imagery accompanied at every step.
First, the regions were separated according to their ownership to the retina or the optical disc. This division occurred as a result of an initial visual assessment showing a clear difference in the system's ability to detect changes in both regions. In each region, changes were divided by sign and classified first by whether they appeared (A), disappeared (D) or had associated changes in the form (F). In a separate set of columns the number of times that these changes were detected was counted as: properly changed, improperly changed, as changes that do not exist or changes not detected.

Thus, depending on the application, the amounts of TP (True Positive), FP (False Positive), TN (True Negative), and FN (False Negative) were specified. Their correspondence with the elements of the tables marked by specialist [12] and other tables generated by them were indicated.

2.7.1. Specification of amounts TP-FP-TN-FN

TP: it is considered true positive to any region, in Table III, marked by the expert as belonging to the column "properly detected". From the detection point of view, it also includes regions belonging to the column "inadequately detected" because, although the detection is inadequate, it exists as such. Figure 2 shows an example of TP which properly recognizes the occurrence of bleeding.

![Figure 2. Example of True Positive (TP).](image)

FP: this includes areas that have been marked by the specialist as belonging to the columns for "improperly detected" and "detected change where there is none?". From the standpoint of detection only, only the regions corresponding to the column mentioned in second place were taken. Figure 3 shows an FP example.

| SIGN | RETINA | A | D | F | adequate | inadequate | not exist | not detected | Total |
|------|--------|---|---|---|----------|------------|-----------|-------------|-------|
| H    | 33     | 5 | 0 | 0 | 1        | 1          | 0         | 0           | 57    |
| E    | 31     | 7 | 6 | 1 | 1        | 0          | 0         | 0           | 14    |
| V    | 13     | 2 | 5 | 2 | 1        | 1          | 0         | 0           | 0     |
| O    | 0      | 0 | 0 | 0 | 0        | 0          | 0         | 0           | 4     |
| R    | 0      | 1 | 0 | 1 | 0        | 0          | 0         | 0           | 1     |
| Total| 77     | 43| 12| 69| 6        | 5          | 3         | 10          | 101   |

| SIGN | Optic Disc | A | D | F | adequate | inadequate | not exist | not detected | Total |
|------|------------|---|---|---|----------|------------|-----------|-------------|-------|
| H    | 3          | 2 | 1 | 0 | 3        | 0          | 0         | 0           | 1     |
| E    | 5          | 0 | 0 | 0 | 3        | 0          | 0         | 0           | 5     |
| V    | 12         | 0 | 7 | 0 | 1        | 4          | 1         | 0           | 1     |
| O    | 4          | 0 | 0 | 0 | 0        | 0          | 4         | 0           | 0     |
| R    | 0          | 0 | 0 | 0 | 0        | 0          | 0         | 0           | 0     |
| Total| 24         | 3 | 7 | 11| 0        | 0          | 13        | 2           | 2     |
Figure 3. Example of False Positive (FP).

TN: For the determination of TN an estimation from the following equation is carried out:

\[ TN=\frac{(AT - ATP - AFP - AFN)}{AAR} \]

Where: AT: area of region of interest of the analyzed image, ATP: total area corresponding to the regions TP; AFP: total area corresponding to the FP regions, AFN: total area corresponding to the regions FN; AAR: average area of TP and FN regions

FN: includes areas that have been marked by the specialist as belonging to the row "changes not detected". Figure 4 shows an example of an FN case where a reddish pigmentation possibly related to a punctate bleeding is not detected.

Figure 4. Example of False Negative (FN).

3. Results
The results are expressed by values of sensitivity and specificity of the system, according to different grouping data criteria as described below. The different ways of grouping the data intended to be used to determine strengths and weaknesses in system performance, depending on the region of the retina under study, the detected signs and changes detection.

Case A: From the general point of view of detection only: Taking into account all the signs, both in the region corresponding to the retina such as the optic nerve, consider all the changes, regardless of whether they are adequate or not.

Case B: From the general point of view of proper detection: Given all the signs, both in the region corresponding to retina as well as the optic nerve, consider only the proper detection and TP.

Case C: From the general point of view of only detection and considering only detected changes on the retina: Given all signs, related only to the region of retina, consider all the changes, regardless of whether they are adequate or not.

Case D: From the general point of view of appropriate detection considering only the changes on the retina: Given all signs, related only to the region of retina, consider only the proper detection and TP.

Case E: From the general point of view of detection only in the optic nerve: Considering all signs, related only to the optic nerve region, regardless of whether or not screening properly.

Case F: From the standpoint of just detecting changes associated with bleeding: Taking only bleeding into account, both in the region corresponding to retina and the optic nerve, consider all the changes, regardless of whether they are adequate or not.

Case G: From the standpoint of proper detection of changes associated with hemorrhage:

SABI 2011 IOP Publishing
Journal of Physics: Conference Series 332 (2011) 012036
doi:10.1088/1742-6596/332/1/012036
Taking only bleeding into account, both in the region corresponding to retina as the optic nerve, consider only the proper detection and TP.

In Table IV results are expressed in different cases in percentage values.

| Data expressed in percentages |
|-----------------------------|
| A  | B  | C  | D  | E  | F  | G  |
|---|---|---|---|---|---|---|
| Sensitivity | 87.62 | 86.95 | 88.09 | 87.34 | 84.61 | 98.24 | 98.21 |
| Specificity | 99.70 | 99.61 | 99.93 | 99.83 | 99.72 | 99.98 | 99.96 |

### 3.1. Observations and analysis of results

The results showed that, for all grouping data criteria, the specificity of the system was above 99.6%.

The high proportion of true negatives with respect to total negative suggests a low probability that the system will detect changes where none exist. This is clearly evidenced in the results if we consider that, of all the regions analyzed, the system erroneously determined as corresponding to changes only 7 regions out of a total of 101 detected.

The overall sensitivity has values close to 87% for both the detection of changes as their appropriate detection. The small gap between the two values indicates that most changes detected by the system are carried out properly.

When data were grouped according to the region in which they are, there is an increase of approximately 5% of sensitivity to changes located in the region corresponding to the retina, with respect to those on the optic nerve. Furthermore, sensitivity is high when working independently with the optic nerve, suggesting the need to work with both regions in different ways, mainly with regard to correction lighting preprocessing stage. Figure 5 shows a case where the system detects changes in the optic nerve, while the expert recognizes such changes as false positives.

![Figure 4](image)

**Figure 4.** (a) First image of the retina. (b) Image of the same eye taken several months later. (c) Output of the system makes changes where there are no such changes (FP).

Independent analysis of the results of changes associated with bleeding with respect to other signs shows a sensitivity and specificity of 98.21% and 99.96% respectively, which corresponds to a sensitivity that is approximately 12% higher than for the general case.

### 3.2. Analysis of processing times

We conducted a study of the time taken for each stage of the workflow in order to recognize those with higher processing time. In turn, we used images with different resolutions in order to observe the variation that this factor produces on the results. Table V shows the times measured in seconds, which was obtained by applying the workflow end, through the generated interface.

We can see from the measurement of processing times that the stages of registration and masking generated the main delay, especially in large images. They have not been optimized so their processing time could be significantly reduced.

The stages which we worked on, tested and adapted showed minimal processing times,
consistent with the purposes of the application. When working with images of standard resolution of 512 x 512, processing times which would make feasible its use in the dynamics of the professional consultation were required.

**Tabla VI:** Duration in seconds of each stage, for pairs of images of size 256 x 256, 512 x 512, 1345 x 1134 and 2048 x 2048 pixels

| Stage                              | 256x256 | 512x512 | 1345x1134 | 2048x2048 |
|------------------------------------|---------|---------|-----------|-----------|
| Registration                       | 0,12    | 0,35    | 1,74      | 5,70      |
| Initial masking                    | 0,21    | 0,51    | 2,88      | 7,81      |
| Ratio G / R for the pair           | 0,018   | 0,062   | 0,36      | 0,89      |
| Statistical normalization for the pair | 0,02   | 0,04   | 0,22      | 0,51      |
| Masking Final                      | 0,12    | 0,47    | 2,75      | 7,35      |
| Simple Difference                  | 0,0004  | 0,0014  | 0,0075    | 0,0044    |
| Thresholds                         | 0,002   | 0,003   | 0,008     | 0,02      |
| Regions and properties             | 0,005   | 0,02    | 0,05      | 0,12      |
| Image Overlay                      | 0,10    | 0,12    | 0,20      | 0,85      |
| Sum Total                          | 1,19    | 1,58    | 8,21      | 23,30     |
| Total Prototype                    | 1,47    | 1,60    | 8,25      | 23,50     |

4. Conclusions and Discussion

In this work we have managed to acquire, standardize and characterize a set of images for further processing, by generation of MGS, selection and adjustment of methodologies and measurement performance of final workflow.

The selection, adaptation and implementation of methodologies successfully resulted in a robust workflow, which is simple, fast, clear and reliable. This is clearly seen on the results of the testing and evaluation of performance, where the sensitivity and specificity exceeded 80%, with processing times lower than 3 seconds for images of 512 x 512 pixels. This permits the inference that the workflow obtained has a high potential to be incorporated into a usable system in the dynamics of ophthalmologists.

The final results were suitable mainly for the determination of changes associated with signs of bleeding. There is a clear need for further testing and performance of the workflow for other signs present on the optical disc.

To adjust the parameters and the selection of methodologies, an MGS demarcated by a single specialist was used. Although this was considered ideal and plausible, we should consider the possibility of demarcation with a representative group of ophthalmologists.

The developed flow could be modified in order to extrapolate these results to other types of signs and fundus images.

Acknowledgements

Special thanks to the collaboration of Dr. Rodrigo Torres (ophthalmologist, Paraná, Entre Ríos), in his capacity as advisor and guide on the subject of monitoring changes in fundus.

5. References

[1] Muhammad-Amri Abdul-Karim et al. 2003 *Automated Tracing and Change Analysis of Angiogenic Vasculature from in vivo multiphoton confocal image time series.* Microvascular
Research 66. Págs. 113-125.
[2] Harihar Narasimha-Iyer et a, l2006. Robust Detection and Classification of Longitudinal Changes in Color Retinal Fundus Images for Monitoring Diabetic Retinopathy. IEEE Transactions On Biomedical Engineering, Vol. 53, Nº 6.
[3] A. Goshatsbsy. 2005 2–D and 3–D Image Registration for Medical, Remote Sensing, and Industrial Applications. Wiley – Interscience,
[4] G. K. Mastopouloes et at. 1999 Automatic Retinal Image Registration Scheme Using Global Optimization Techniques. IEEE-Trans. Inf. Technol. Biomed., Vol. 3, Nº 1, págs. 47-60.
[5] N. Ritter et al. 1999 Registration of Stereo and Temporal Images of the Retina. IEEE Trans. Med. Imag., Vol. 18, Nº 5, págs. 404-418.
[6] C.V. Stewart et al. 2003 The Dual-bootstrap Iterative Closest Point Algorithm with Application to Retinal Image Registration. IEEE Trans. Med Imag., Vol. 22, Nº 11, págs. 1379–1394.
[7] A. Can et al. 2002 A Feature-based, Robust, Hierarchical Algorithm for Registratering Pairs of Images of the Curved Human Retina. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 24, Nº 3, págs 347-364.
[8] R.C. Gonzales y R.E. Woods. 1992 Digital Image Processing. Reading, MA: Addison-Wesley.,
[9] B.H. Brinkmann et al. 1998 Optimized Homomorphic Unsharp Masking por MR Grayscale Inhomogenity Correction. IEEE Trans. Med. Imag., Vol. 17, Nº 2, págs. 161-171.
[10] R. Lililestrand. 1972 Techniques for Change Detection. IEEE Trans. Comput., Vol. 21, Nº 7, págs. 655–659.
[11] X. Dai y S. Khorram, 1998 The Effects of Image Misregistration on the Accuracy of Remotely Sensed Change Detection. IEEE Trans. Geosci. Remote Sens., Vol. 36, Nº 5, págs. 1566–1577.
[12] R. J. Radke et al. 2005 Image Change Detection Algorithms: A Systematic Survey. IEEE Transactions on Image Processing, Vol. 14, Nº 3.
[13] Li Chen. 2006 Feature-Based Retinal Image Registration Using Bifurcation Structures.
[14] A. Elgammal et al. 2002 Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance. Proc. IEEE, Vol. 90, Nº 7, págs. 1151–1163.
[15] N. M. Salem y A.K. Nandi. 2007 Novel and Adaptive Contribution of the Red Channel in Pre-Processing of Colour Fundus Image”. Journal of the Franklin Institute.
[16] L. P. Rosin et al. 2003 Evaluation of Global Image Thresholding for Change Detection. Department of Computer Science, Cardiff University, UK.
[17] L. P. Rosin. 2002 Thresholding for Change Detection. Comput. Vis. Image Understanding, Vol. 86, Nº 2, págs. 79–95.
[18] P. Smits y A. Annoni. 2000 Toward Specification-Driven Change Detection. IEEE Trans. Geosci. Remote Sens., Vol. 38, Nº 3, págs. 1484–1488.
[19] R. M. Haralick y L. G. Shapiro. 1992 Computer and Robot Vision. Reading, MA: Addison-Wesley, Vol. 1.
[20] E. Stringa. 2000 Morphological Change Detection Algorithms for Surveillance Applications. Proc. Brit. Machine Vision Conf.
[21] Tarkan Sevilmis et al. 2008 Automatic Detection of Salient Objects and Spatial Relations in Videos for a Video Database System. Image and Vision Computing, Vol. 26, págs. 1384–1396.
[22] 22.M. Tedder y C. Jin-Chung. . 2004 Autonomous Robot Vision Software Design Using Matlab Toolboxes. Lawrence Technological University.