

Original article

TeleOphta: Machine learning and image processing methods for teleophthalmology

E. Decenci re^a, G. Cazuguel^{d,f}, X. Zhang^{a,*}, G. Thibault^a, J.-C. Klein^a, F. Meyer^a,
B. Marcotegui^a, G. Quellec^d, M. Lamard^{d,g}, R. Danno^e, D. Elie^e, P. Massin^b, Z. Viktor^b,
A. Erginay^b, B. La y^e, A. Chabouis^c

^a Center for mathematical morphology, systems and mathematics department, MINES ParisTech, 35, rue St-Honor , 77300 Fontainebleau, France

^b Service d'ophtalmologie, h pital Lariboisi re, AP-HP, 2, rue Ambroise-Par , 75475 Paris cedex 10, France

^c Direction de la politique m dicale, parcours des patients et organisations m dicales innovantes – t l m decine, Assistance publique–H pitaux de Paris, 3, avenue Victoria, 75184 Paris cedex 04, France

^d Inserm UMR 1101 LaTIM, b timent I, CHRU Morvan, 29200 Brest, France

^e ADCIS, 3, rue Martin-Luther-King, 14280 Saint-Contest, France

^f T l com Bretagne, Institut Mines-T l com, ITI Department, 29200 Brest, France

^g Universit  de Brest, Inserm UMR 1101 LaTIM, SFR ScInBioS, 29200 Brest, France

Received 21 January 2013; received in revised form 21 January 2013; accepted 21 January 2013

Available online 15 March 2013

Abstract

A complete prototype for the automatic detection of normal examinations on a teleophthalmology network for diabetic retinopathy screening is presented. The system combines pathological pattern mining methods, with specific lesion detection methods, to extract information from the images. This information, plus patient and other contextual data, is used by a classifier to compute an abnormality risk. Such a system should reduce the burden on readers on teleophthalmology networks.

  2013 Elsevier Masson SAS. All rights reserved.

1. Introduction

Diabetic retinopathy is one of the major causes of visual impairment and blindness in the working population. Early detection of diabetic retinopathy improves healing chances or helps stopping or slowing down its progress. In order to achieve early detection, national and international guidelines recommend annual eye screening for all diabetic patients. However, annual fundus examination is not performed sufficiently. One of the factors which explains this situation is the increasing number of diabetic patients (a worldwide phenomenon) and the decreasing number of ophthalmologists (at least in France). A screening program by telemedicine can improve the regular annual fundus examination of diabetic patients.

The first teleophthalmology networks, such as OPHDIAT [1,2], are devoted to the early detection of retinal retinopathy.

Currently, the resulting fundus photographs are manually graded by ophthalmologists. However, ophthalmologists are already overwhelmed with work, and their numbers are decreasing. Given the increasing incidence of diabetes [3], new screening centers are expected to join the existing telemedicine networks, and new networks should be created. Therefore, this situation should not improve in a near future. So the interpretation of fundus photographs will represent a huge burden. It should be noted that, among the patients screened every year, about 75% have no diabetic retinopathy and only about 26% were referred to an ophthalmologist for either diabetic retinopathy or other pathologies [1]. A method to automatically detect healthy cases in telemedical networks would foster the development and deployment of these networks.

It was in this context that the partners of TeleOphta proposed to develop a system to classify the examinations acquired in a teleophthalmology network, in order to classify them into one of two categories: “normal” or “to be referred to a specialist”. These examinations contain color eye fundus images, as well as contextual data, including patient data. The French National Research

* Corresponding author.

E-mail address: xiwei.zh@gmail.com (X. Zhang).

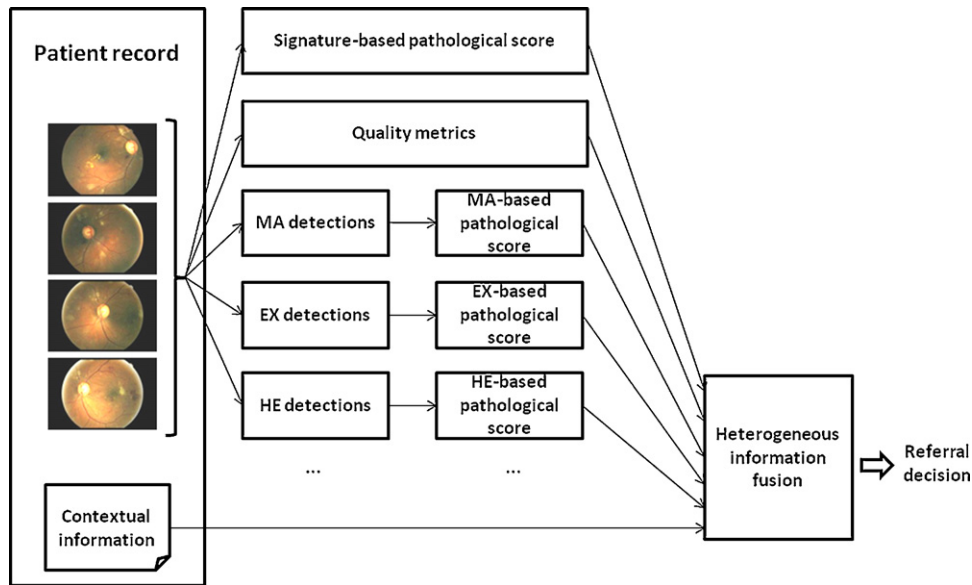


Fig. 1. Illustration of the TeleOphta strategy. Patient record classification workflow.

Agency (ANR), through its TECSAN program, accepted to fund the project in 2009.

The TeleOphta project has benefited from a large amount of data extracted from OPHDIAT. It has developed image processing and machine learning methods to classify each examination into “normal” or “to be referred”. The main pathology that is taken into account is diabetic retinopathy, but the project objectives include the detection of other retinal pathologies.

The objective of this paper is to give a general overview of TeleOphta. It is not the authors’ intention to delve into technical details. Moreover, even if the project developments are in their last phase, some are not completely finished or validated. Readers interested by the details should consult previously published articles, and stay tuned for future publications by the TeleOphta consortium.

2. General strategy

The development of image processing methods for the analysis of eye fundus images has been a very active research domain for the last 15 years. Naturally, the first studies concerned relatively small and homogeneous databases. However, in order to bridge the gap between academic solutions and clinical applications, especially in the framework of teleophthalmology, an eye fundus analysis method should be able to deal with a large variety of images. Indeed, image quality and size, acquisition conditions, among other characteristics, can change greatly from one examination to another. One of the major challenges TeleOphta had to deal with was this image heterogeneity.

However, images are not the only source of information taken into account by the specialist to produce a diagnostic. Contextual information, as patient age or diabetes history, is also considered. Therefore, in the framework of TeleOphta, machine learning methods have been used to fusion information obtained from the images with contextual information. This was the second major challenge the project had to solve.

Finally, even if TeleOphta has been designed to process data acquired on a network devoted to diabetic retinopathy screening, the proposed solution should be able to detect all abnormal examinations, including other retinal pathologies.

In order to rise to these challenges, the TeleOphta partners proposed to combine generic image signatures, as well as specific lesion detection methods, with classification methods. In order to develop and optimize these methods, a large amount of anonymized data was extracted from the OPHDIAT network. This general strategy is illustrated in Fig. 1.

3. Databases

The e-ophtha database results from the anonymized extraction of the examinations gathered during years 2008 and 2009 through the OPHDIAT network. This represents 25,702 examinations, each one containing at least four images (two per eye) as well as contextual data, as patient age, number of years of diabetic retinopathy, etc. Finally, each examination is accompanied by the text annotations of the OPHDIAT reader: diabetic retinopathy level for each eye, as well as “course to follow” (such as one year follow-up examination, or referral to an ophthalmologist). It should be noted that the images are compressed with the JPEG standard, for practical reasons linked to network bandwidth. The whole database weights 50 GB.

Five hundred examinations were chosen from e-ophtha in order to proceed to two supplementary readings. Thus, three expert opinions will be available for each examination on this database, called e-ophtha DR (“double-reading”).

In order to develop the lesion detection methods, lesions were manually outlined on some images by an ophthalmologist using software developed by ADCIS (Annex). These annotations were afterwards checked by a second ophthalmologist.

Two such databases should be made available before the end of the project. The first, e-ophtha EX, contains 47 images with 12,278 exudates, as well as 35 healthy images. Several images

among the healthy ones contain structures which can easily mislead exudate detection methods, such as reflections, and optical artifacts. Some image examples are shown in Fig. 2. The second of these databases, called e-ophtha MA, contains 148 images with 1306 microaneurysms, as well as 233 healthy images. A third lesions database, containing hemorrhages, should be available shortly afterwards.

The authors believe that these databases constitute a major contribution for the research community. They gather a huge amount of precious information, useful for the development of automatic screening methods. We will see for instance in the following section how e-ophtha is used to design pathological pattern mining methods.

4. Mining pathological patterns in images

We present in this section a general solution to roughly detect the signs of DR, and of other retinal pathologies. The proposed solution relies on wavelet-based image characterizations developed in previous works [4].

Each image in a patient record is divided into patches and a vector of image characterizations, called signature, is extracted from each patch. A machine learning algorithm was designed to recognize those signatures that only appear in pathological patient records [5]. This algorithm relies on the multiple-instance learning paradigm. In order to detect various pathological patterns, several sizes of patches are used simultaneously [6]. A global pathological index is then derived for the patient record as a whole: it combines local pathological scores computed in image patches individually [7]. The algorithm is summarized in Fig. 3. In order to push the classification performance further, the shape of the wavelet filters used to extract image characterizations is tuned by a genetic algorithm [8].

Note that this signature-based detector is not supervised by manual segmentations. Instead, it is supervised by the decision attached to patient records as a whole: whether or not the patient should be referred to an ophthalmologist. A subset containing 800 records from e-ophtha was used for training.

This generic approach allows automatically detecting pathological patterns in images. Therefore, we can expect that pathologies which are conveniently represented in the database will be detected with a high probability. *Per se*, this approach already provides meaningful information to the patient records classifier, resulting in an interesting trade-off between sensitivity and specificity. However, from a scientific and application point of view, it was interesting to see if adding specific information, brought by more classical lesion segmentation methods, would improve the performance of the system.

5. Specific lesion detection methods

The final objective of the image processing methods developed in the framework of TeleOphta is to extract meaningful data from the images in order to provide information to the patient records classifier. We give below a brief description of the main methods.

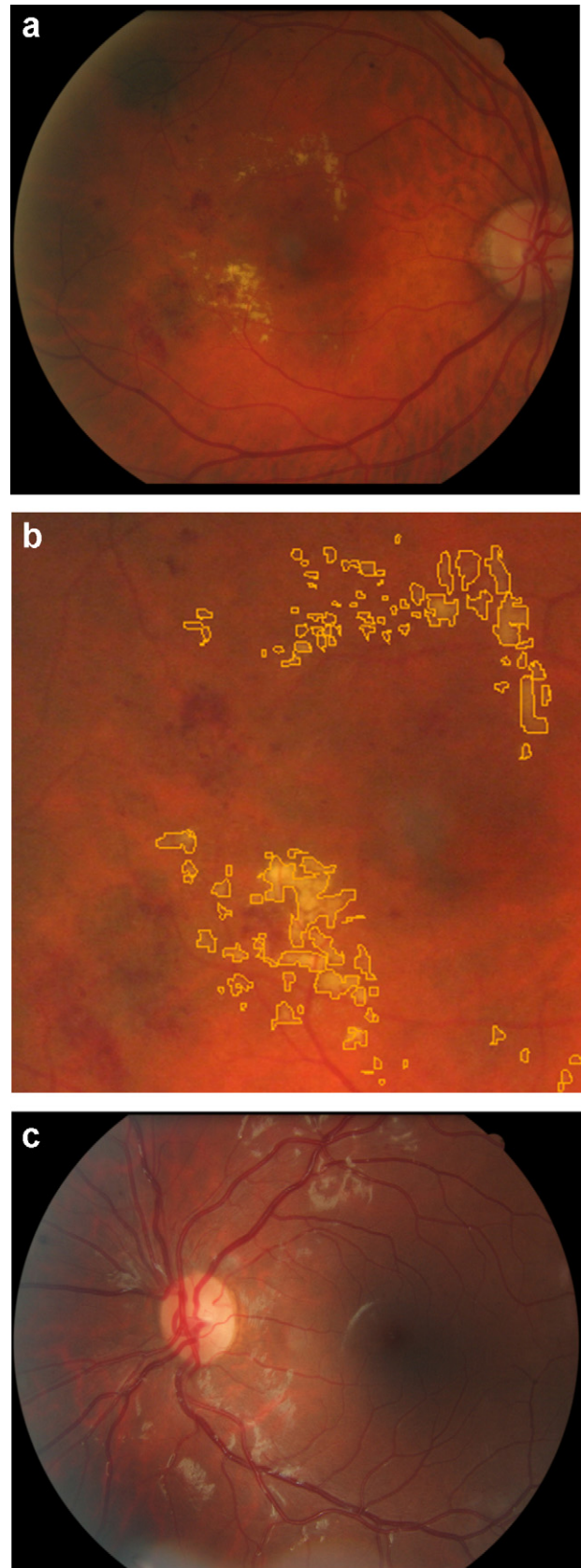


Fig. 2. Image examples: a: image from e-ophtha EX containing exudates; b: detail of (a), showing manually drawn exudate contours; c: healthy image from e-ophtha EX – notice reflections on the retina, which can mislead exudates detection methods.

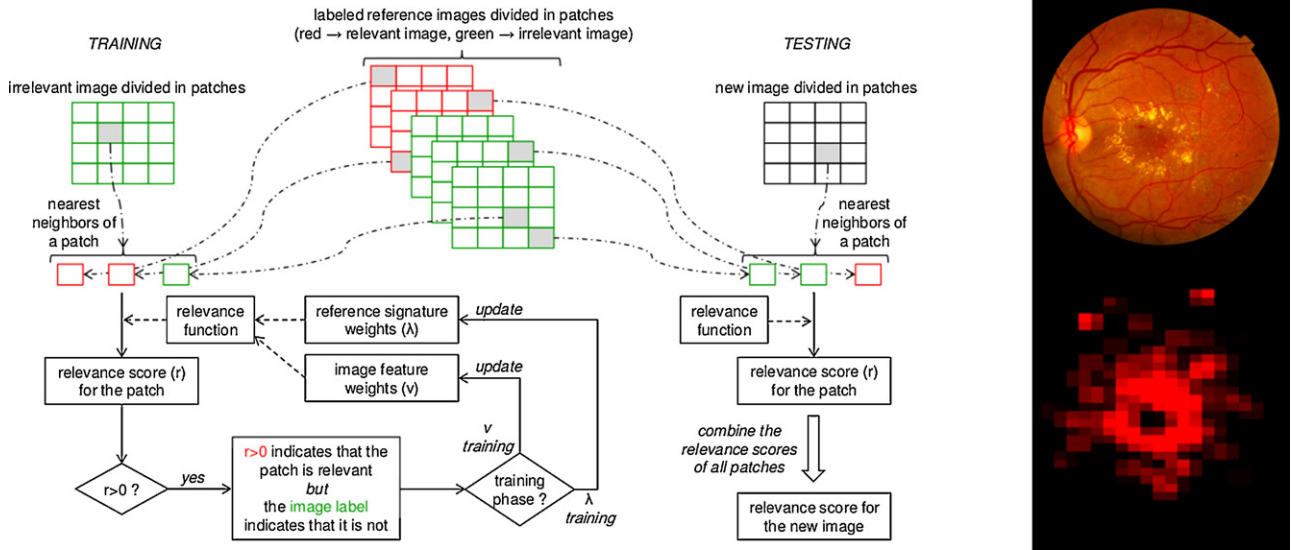


Fig. 3. Mining pathological patterns in images.

5.1. Preprocessing

Preprocessing is particularly important for our application, due to the clinical nature of the database.

Given the heterogeneity of the images we have to deal with, a specific spatial normalization method has been developed [9,10]. This method uses the field of view (FOV) of the fundus photographs as size invariant (a hypothesis which has proven to be robust enough thanks to the fact that all the images are acquired with the same field of view angle, i.e. 40°). Using this size invariant, all parameters related to size can automatically be defined, once they have been chosen for a given image resolution. In the diagram given in Fig. 4, the output of this module is called “pixel size”. This must be understood as a *relative* pixel size.

Finally, border reflections are segmented. These reflections appear as bright, moon-shaped regions along one side of the field of view. They can perturb the detection of structures such as the optic disk. This segmentation has two steps. First, the residue between the original blue channel and its response to a large mean filter leaves only the main bright regions. Then a reconstruction from the border of the FOV gives the reflection mask. As shown in Fig. 4, this module takes as input not only the original color image, but also the pixel size obtained from a previous step. It should be noted that taking into account this kind of artifact is, to the limit of the authors’ knowledge, an original contribution. Indeed, until now images considered in the literature are relatively “clean”, without optical or other artifacts, which are found in practice, and can be a serious source of false positives or negatives.

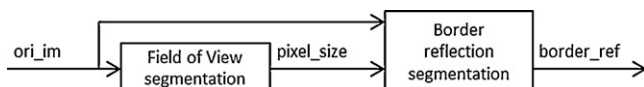


Fig. 4. Preprocessing steps diagram; ori corresponds to the original color image of the retina.

Once the preprocessing steps are finished, we can move to the detection of anatomic structures.

5.2. Anatomic structures detection

We need to detect the main anatomical structures of the retina in order to improve the detection of lesions. As shown in Fig. 5, we first segment the main vessels, then we detect the optic disk, and finally the macula.

For our current purposes, we only need a rough segmentation of the retinal vessels. We do not (yet) bother about small vessels. The resulting vessels mask might even contain lesions, such as hemorrhages and microaneurysms. Later, the vessels mask will be analyzed to extract from it those structures. From the time being, a morphological alternate sequential filter is used to erase vessels, as well as other structures from the original image. The positive residue of this operator (i.e. the positive part of $ASF(f)-f$, where f is the green channel of the original color image, and ASF is the alternate sequential filter) is thresholded to obtain the vessels, along with noise. Then a length opening removes small structures and keeps the mask we are looking for.

The Optic Disc (OD) is the brightest structure in normal eye fundus images. Besides its high intensity, the OD is also the main convergence point of the retinal vessel network. The detection starts with a threshold to get some large bright structures as OD candidates. The previously computed vessel mask is used to select the exact position and to check the presence of the OD. The skeleton, width, density and orientation of the vessel mask are computed. The OD region is supposed to contain the highest density of wide vertical vessels.

In contrast, the macula appears most of the time as a dark region. Its distance to the OD is approximately known. In order to combine these two hypotheses, we threshold the image to detect dark regions, and then keep only those whose distance to the OD verify the second hypothesis.

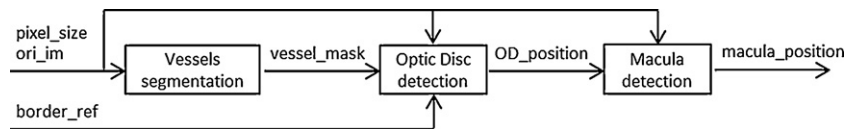


Fig. 5. Anatomic structures segmentation and detection.

5.3. Lesions detection

A new exudates detection method has been developed for heterogeneous databases [11]. Its main originality lies probably on the fact that it explicitly detects artifacts or structures which can be erroneously considered as exudates. The OD and border reflections have already been detected in the preprocessing phase. In clinical practice, reflections in the middle of the vessels and along the vessels are very common. Bright artifacts caused by camera lens should also be considered. We have shown that the saturation channel of the image allows detecting these artifacts. The morphological ultimate opening [12] is then combined with a local variance threshold to extract exudate candidates. Finally, about 30 characteristics are extracted from each candidate to train a random forest model [13]. The classifier associates to each candidate a risk value, which can be interpreted as the probability of the candidate being a real exudates.

Detecting microaneurysms on the e-ophta MA database has proven to be extremely difficult. Based on previous approaches [14,15], we have simplified the candidate detection step and shifted most of the work to the feature extraction step, so that a trained classifier makes the final decision. The candidates are obtained with the same alternate sequential filter as the vessel segmentation. But this time, only small structures are kept. Features are extracted from each candidate, by means of a max-tree [16] decomposition, including local, geometrical and contextual properties. As for the exudates, a random forest classifier is used to compute the final risk associated to each candidate.

Hemorrhages detection is a very difficult challenge, seldom treated in the literature. The main difficulties come from their variable shape, size and contrast. They can be almost as small as microaneurysms or as large as the OD. Flame-hemorrhages have the same shape as vessels. We have proposed a new method, which begins by roughly segmenting the vessel network and other dark structures. This new vessel mask contains vessels as well as hemorrhages. Vessels width and orientation are computed on the skeleton of the vessel mask. Each skeleton branch (i.e. skeleton section between two consecutive skeleton bifurcations) is labeled, while all bifurcation and end points are detected. Thus, using the contextual and geometrical properties of the vessel network, abnormal points are detected as hemorrhage candidates. The final risk is again obtained with a random forest classifier (Fig. 6).

It should be noted that in the framework of TeleOphta previously existing microaneurysm and exudates detection methods have also been used. Indeed, as stated before, the system is open and modular, and therefore supplementary lesion detection methods can be easily added.

The lesion detection methods presented until now focus on the most current lesions of diabetic retinopathy. What

about other lesions, such as cotton-wool spots, neo-vessels, and intra-retinal microvascular abnormalities, which are often very important from the diagnosis point of view? Given the generic pathological pattern mining approach described in the precedent section, we believe that most of these lesions will be taken into account in the final decision. Indeed, we will see that the current version of the TeleOphta system already provides very interesting results. However, adding more specific detection methods is an improvement path that we are going to explore in the future.

6. Classifying patient records

Now that the visual content of images has been characterized, we present how image characterizations are combined with contextual data (age, weight, diabetes type, etc.) in order to decide whether or not a patient should be referred to an ophthalmologist. This classification problem has two main challenges. First, we need to process a varying number of lesion detections per patient record. Second, we need to process sparse contextual data. The workflow is summarized in Fig. 1.

6.1. Lesion-based pathological score

To address the first challenge, we first compute a single pathological score per patient record and per lesion detector. In that purpose, the joint cumulative distribution function (CDF) of the lesion probabilities and of the lesion sizes is built for each patient record. This CDF is then mapped to a single pathological score using a linear discriminant analysis [17]. The mapping process is tuned to maximize classification performance in the training subset of e-ophta.

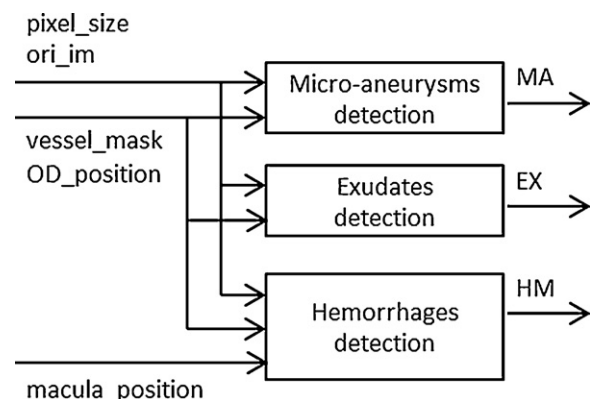


Fig. 6. Lesions detection.

6.2. Heterogeneous information fusion

The second step consists in classifying a sparse vector of heterogeneous descriptors: one pathological score per lesion detector, one signature-based pathological score, six quality metrics, up to nine demographic information fields and up to 18 diabetes-related information fields. Solutions based on decision trees and the Dezert-Smarandache theory have been proposed in previous works [18]. A new solution based on Apriori, the most popular algorithm for association rule mining [19], was proposed in this study. Rules associating a subset of descriptors with a referral decision (e.g. $\{70 \leq \text{age} < 75, \text{diabetes_type} = \text{NIDDM}, \text{MA_score} > 0.75\} \rightarrow \text{refer_to_ophthalmologist}$) are mined in the training subset of e-optha. Their sensitivity and specificity are measured in the training subset. To classify a new patient record, all the relevant association rules are selected. Then, a referral decision is inferred based on the sensitivity and specificity of the selected rules. This solution is particularly well suited to sparse data.

7. Integration and preliminary results

Algorithms developed by both research groups were delivered to ADCIS for integration into the same software environment, and for validation on 1,000 examinations. Algorithms originally developed in proprietary languages and environments (i.e. C++, Python, Java) were rewritten by the ADCIS engineering group to make them compatible with the company environment, and to be reused in future projects.

The training of information fusion rules was performed on half of the e-optha database, i.e. 12,500 examinations. The test was performed on the other half of the database. The system gives as first output a “To be referred” probability. Thresholding this value leads to a referral decision. Then a ROC curve was computed to evaluate the automatic grading system (Fig. 7), with OPHDIAT ophthalmologists decision (to be referred to an ophthalmologist or not) as ground truth. Two micronaneurysm detection methods, as well as two exudate detection methods were used, on top of patient data and image signatures.

The ROC curve shows that for the same sensitivity as the human expert (blue star, sensitivity of 80,9%), our system has a lower specificity (70% instead of 81,5%). But it means that it is possible to remove 70% of “Not to be referred” patients from the patients records sent to the OPHDIAT network ophthalmologists. Knowing that they represent 74% of the cases, the proposed system will be able, in this example, to remove about 50% of records which have to be actually examined in the OPHDIAT network. With a sensitivity of 90%, this value goes down to 35% of the total records, but it is still very interesting: it allows to save 35% of ophthalmologist time.

Nevertheless, although this ROC curve shows the ability of the system to classify exams in two groups (normal and to be referred to a specialist), an additional validation will be performed by the medical partners on 1,000 exams of year 2012 to show the effectiveness of the developed method. This validation is currently under progress.

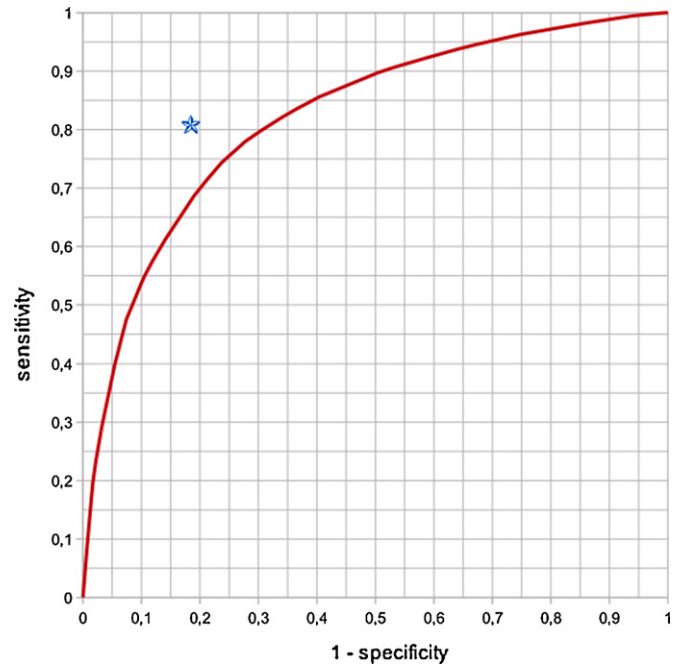


Fig. 7. ROC curve obtained on the e-optha database (in red). Patient data, image signature, two exudate detection methods, and two microaneurysms detection methods have been used to feed the patient records classifier. The blue star, in comparison, shows the sensitivity (80.9%) and specificity (81.5%) obtained by a human expert on a subset of 500 exams extracted randomly from the e-optha database. Note that this human expert was not exactly in the same conditions of interpretation than OPHDIAT experts.

Concerning future clinical trials, the medical partners will choose the system sensitivity, taking into account double-reading results on the 500 cases of the e-optha DR database (in construction).

8. Conclusion

The TeleOphta project has developed a novel strategy to relieve the burden on teleophthalmology networks, through the combination of an original pathological pattern mining method, with new specific lesion detection methods. The final classification takes also into account contextual information.

One of the main results of the project consists in the databases. Given their clinical nature, they will bring valuable data to the research community, and should contribute to the improvement of ophthalmology screening methods.

Concerning image processing methods, a strong emphasis has been put on the development of robust methods, in order to tackle the challenge of the heterogeneity of the databases. The resulting exudates and microaneurysms detection methods are being validated on the project databases. The detection of hemorrhages represents an even greater challenge, given their potential resemblance to blood vessels.

The system is open, and its performance could be improved by adding new specific lesion detection methods, especially for neo-vessels, intra-retinal microvascular abnormalities and other abnormalities of the vessel network. A precise vessel segmentation method, able to detect even small vessels on less-than-ideal quality images will be necessary for these tasks.

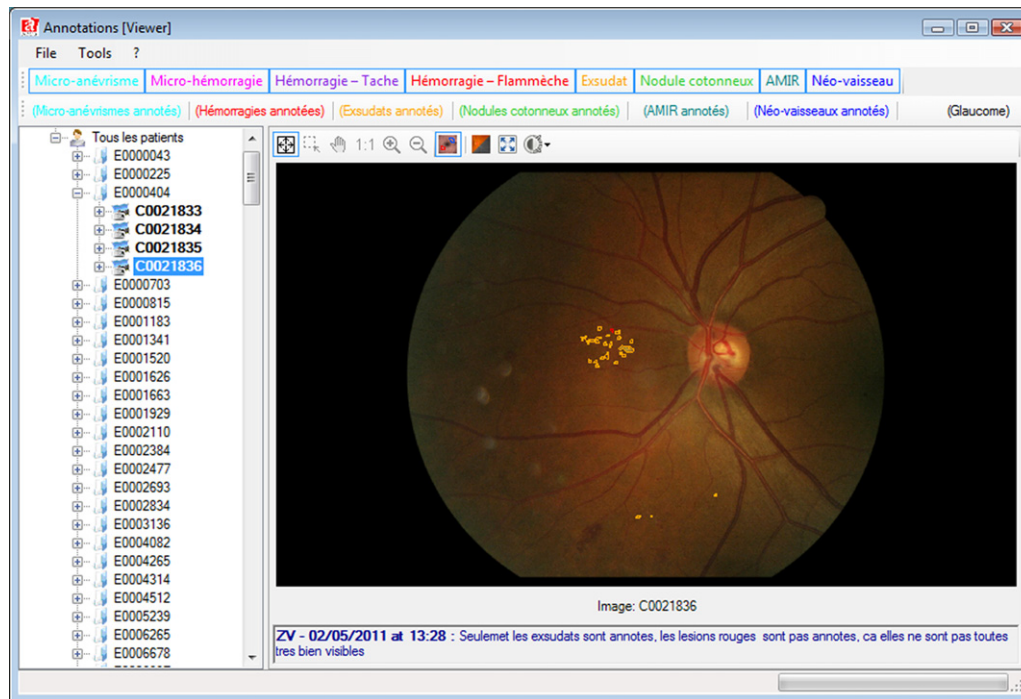


Fig. 8. Overview of the TeleOphta annotation software, in Viewer mode. All images, organized in folders, appear in the left-hand pane. The selected image is displayed in the right-hand pane, with annotations in overlay (several exudates appear in orange).

The different methods developed in the framework of TeleOphta have been integrated into a single system. The first results show that the performance of this system should lead to an interesting trade-off between sensitivity and specificity. Based on the obtained ROC curve, it will be easy to tune the final system for clinical evaluation.

Finally, the proposed system will allow saving ophthalmologist time in screening networks, and thus will help rising to new population health challenges.

Acknowledgements

This work has been supported by the French National Research Agency (ANR) through TECSAN programme (ANR-09-TECS-015).

Annex. Annotation software

An image annotation software was especially designed to facilitate the communication between ophthalmologists and image processing researchers, even if they are remotely located. Users of this specific software are divided in two categories.

On the one hand, the Grader (an ophthalmologist) outlines objects of interest or marks them by a cross using the mouse pointer, and then assigns the considered objects to a class from a pull-down list of predefined classes. Predefined classes consisted of: microaneurysms, dot-hemorrhages (marked by a cross), blot-hemorrhages, flame-hemorrhages, exudates, cotton-wool spots, intra-retinal microvascular abnormalities and neo-vessels (outlined). Optionally, comments may be added to the whole image. A set of tools are available to help in the diagnostic, and to

better see small details in the image of interest: zoom in/out, display/undisplay existing annotations, display green image band only, gamma correction.

Once an annotation or comment is saved by the grader, it is automatically uploaded to a remote server. If different graders annotate the same image, each grader can only see her/his annotations, to avoid any bias. No programming or image processing skills are needed on Graders side. During the course of the project, the annotation software was used by a grader from the AP-HP in Paris.

On the other hand, Viewers (image processing researchers) can see the annotations made by all Graders. Synchronization of the software with the remote database is handled automatically. Once the database is synchronized, Viewers can export annotations to TIFF files, so as to easily compare them with detections produced by their algorithms.

To quickly adapt to the diversity of possible applications, the annotation software developed for TeleOphta gave birth to Image Annotator, a fully-configurable software that is now delivered by ADCIS to its customer community (Fig. 8).

References

- [1] Massin P, Chabouis A, Erginay A, Viens-Bitker C, Lecleire-Collet A, Meas T, et al. OPHDIAT: a telemedical network screening system for diabetic retinopathy in the Île-de-France. *Diabetes Metab* 2008;34(3): 227–34.
- [2] Erginay A, Chabouis A, Viens-Bitker C, Robert N, Lecleire-Collet A, Massin P. OPHDIAT: quality-assurance programme plan and performance of the network. *Diabetes Metab* 2008;34(3):235–42.
- [3] Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004;27(5):1047–53.

- [4] Quellec G, Lamard M, Josselin PM, Cazuguel G, Cochener B, Roux C. Optimal wavelet transform for the detection of microaneurysms in retina photographs. *IEEE Trans Med Imaging* 2008;27:1230–41.
- [5] Quellec G, Lamard M, Cazuguel G, Abràmoff MD, Cochener B, Roux C. Weakly supervised classification of medical images. In: *Proc IEEE ISBI*. 2012, p. 110–3.
- [6] Quellec G, Lamard M, Cochener B, Decencière E, Lay B, Massin P, Roux C, Cazuguel G. A General framework for detecting diabetic retinopathy lesions in eye fundus images. In: *Proc Int Conf IEEE CBMS*. 2012. p. 1–6.
- [7] Quellec G, Lamard M, Abràmoff MD, Decencière E, Lay B, Erginay A, et al. A multiple-instance learning framework for diabetic retinopathy screening. *Med Image Anal* 2012;16(6):1228–40.
- [8] Quellec G, Lamard M, Cazuguel G, Cochener B, Roux C. Wavelet Optimization for content-based image retrieval in medical databases. *Med Image Anal* 2010;14:227–41.
- [9] Zhang X, Thibault G, Decencière E, Guellec G, Gazuguel G, Erginay A, et al. Spatial normalization of eye fundus images. In: *IEEE ISBI*. 2012.
- [10] Zhang X, Thibault G, Decencière E. Procédé de normalisation d'échelle d'images ophtalmologiques. 12 53929 (Patent) 2012.
- [11] Zhang X, Thibault G, Decencière E, Guellec G, Danno R, Lay B, et al. Automatic Detection of Color Retinal Images. Fort Lauderdale, Florida, USA: ARVO; 2012.
- [12] Beucher S. Numerical residues. In: *Proc ISMM*. 2005, p. 23–32.
- [13] Breiman L. Random forests. *Mach Learn* 2001;45(1):5–32.
- [14] Walter T, Massin P, Erginay A, Ordonez R, Jeulin C, Klein JC. Automatic detection of microaneurysms in color fundus images. *Med Image Anal* 2007;11:555–66.
- [15] Zhang X, Thibault G, Decencière E. Application of the morphological ultimate opening to the detection of microaneurysms on eye fundus images from a clinical database. ICS, Oct 2011, Beijing, China.
- [16] Salembier P, Oliveras A, Garrido L. Antiextensive connected operators for image and sequence processing. *IEEE Trans Image Process* 1998;7(4):555–70.
- [17] Quellec G, Lamard M, Cochener B, Droueche Z, Lay B, Chabouis A, et al. Studying disagreements among retinal experts through image analysis. In: *Proc Int Conf IEEE EMBS*. 2012, p. 5959–62.
- [18] Quellec G, Lamard M, Cazuguel G, Bekri L, Daccache W, Roux C, et al. Automated assessment of diabetic retinopathy severity using content-based image retrieval in multimodal fundus photographs. *Invest Ophthalmol Vis Sci* 2011;52:8342–8.
- [19] Agrawal R, Srikant R. Fast algorithms for mining association rules in large databases. In: *Proc Int Conf VLDB*. 1994, p. 487–99.