

CANDIDATURE A LA QUALIFICATION AUX FONCTIONS DE MAITRE DE CONFERENCES

Section CNU: 61^{ème}

Joan Massich

16/12/2016

Etat civil

| | |
|------------------------------|--|
| Nom : | Massich |
| Prénom : | Joan |
| Date et lieu de naissance : | 15 Mars 1984 à Palafrugell (Espagne) |
| Nationalité : | Espagnol |
| Situation de famille : | Célibataire |
| Coordonnées professionnelles | Le2i - UMR CNRS 6306 12 Rue de la Fonderie 71200 Le Creusot +33 6 01 20 16 68 mailsik@gmail.com |
| Coordonnées personnelles : | Joan Massich 30bis Rue du bois La Chapelle du bas 89290 Venoy +33 6 01 20 16 68 mailsik@gmail.com |

Résumé

| | | | |
|-----------------|---|----------------|-------------------------------|
| Publications : | Revues : | 3 | 3 en auteur correspondant |
| | Conférences | 23 | 6 en 1 ^{er} auteur |
| | Rapport technique | 2 | |
| Enseignements : | Architecture des systèmes | 20h | DUT IQ 1 ^{ère} année |
| | Conception Orientée Object | 72h | DUT IQ 1 ^{ère} année |
| | Programmation Web Orienté Client | 44h | DUT IQ 2 ^{ème} année |
| | Programmation mobile | 28h | DUT IQ 2 ^{ème} année |
| | Total équivalent TD | 117.3 h | |

Mots-clés : Medical Image Analysis, Pattern Recognition, Machine Learning, Computer Aided Diagnosis, Computer Vision, Segmentation, Statistical Models, Breast Cancer

Joan Massich

Developer, Research fellow

12 rue de la Fonderie
71200 Le Creusot
+33 (0) 601 201 668
mailsik@gmail.com
www.massich.cat
 massich
 massich
 jmassich
Born 15 March 1984



This is how I did it Anton. I never saved anything for the swim back. (Gattaca, 1997)

Experience

- Dec. 2015 – Sep. 2016 **Attaché temporaire d'enseignement et de recherche (ATER)**, Le2i - Laboratoire Electronique, Informatique et Image UMR CNRS 6306 at IUT Dijon, Dijon
- Dec. 2013 – Dec. 2015 **Postdoctoral fellow**, Le2i - Laboratoire Electronique, Informatique et Image UMR CNRS 6306 at IUT Le Creusot, Le Creusot
- Oct. 2009 – Dec. 2013 **Assistant de recherche**, ViCOROB - Computer Vision and Robotics group at Universitat de Girona, Girona
- June 2009 – Oct. 2009 **Assistant de recherche**, Applied Vision Laboratory (AVL) at Texas Tech University (TTU), Lubbock
- Dec. 2003 – June 2009 **Assistant de recherche**, ViCOROB - Computer Vision and Robotics group at Universitat de Girona, Girona

Education

- Oct. 2009 – Dec. 2013 **Thèse doctorat**
Universitat de Girona, Girona, Catalonia.
Université de Bourgogne, Le Creusot, France. title: Deformable object segmentation in ultra-sound images.
grade: Unanimity excellent with honors.
- Sept. 2007 – June 2009 **Erasmus Mundus Master Course on Computer Vision and Robotics (VIBOT)**
Heriot-Watt University, Edinburgh, Scotland.
Universitat de Girona, Girona, Catalonia.
Université de Bourgogne, Le Creusot, France.
grade: 8.1 over 10
- Sept. 2005 – Sept. 2007 **Ingénierie informatique**,
Universitat de Girona, Girona, Catalonia
grade: 7.7 over 10
- Sept. 2002 – Sept. 2005 **Ingénierie technique dans les systèmes informatiques**,
Universitat de Girona, Girona, Catalonia
grade: 6.9 over 10

Thèse de doctorat

Titre *Deformable object segmentation in ultra-sound images.*

Supervisors Joan Martí and Prof. Fabrice Meriaudeau

Description Sa thèse est consacrée à la segmentation automatique des lésions mammaires dans les images échographiques car cette tâche est essentielle pour le développement de systèmes robustes de diagnostic assisté par ordinateur (CAD) appliqués à cet organe et à cette modalité d'image. La stratégie de segmentation proposée divise les images en régions significatives appelées super-pixels et les étiquette en utilisant un cadre de minimisation qui prend en compte la formation et la régularisation.

Technical Skills

Programming Languages Python, C++, Matlab

Technologies Git, Travis, CMake, L^AT_EX, Vim, Emacs

Operating System Linux (Gentoo)

Languages

Catalan IRL 5

langue maternelle

Spanish IRL 5

langue maternelle

English IRL 4-5

French IRL 3-4

Awards

Feb. 2016

- Best student paper award for, Rastgoo et. al. "Classification of melanoma lesions using sparse coded features and random forests", Rome

Aug. 2006

- Member of the ViCOROB-UDG team honored as the winner of the 2006 SAUC-E (Student Autonomous Underwater Challenge Europe) competition, London

1 Activités pédagogiques

1.1 Résumé des enseignements

Le tableau 1.1 résume les enseignements que j'ai réalisé ces dernières années. Pour plus de détails concernant ces enseignements, reportez vous à la section 1.2.

| Année | Enseignement | Niveau | Volume | | |
|--------------|----------------------------------|-------------------------------|------------|-------------|-----|
| | | | CM | TD | TP |
| 2015/2016 | Architecture des systèmes | DUT IQ 1 ^{ère} année | | | 20h |
| | Conception Orientée Object | DUT IQ 1 ^{ère} année | | | 72h |
| | Programmation Web Orienté Client | DUT IQ 2 ^{ème} année | 12h | | 32h |
| | Programmation mobile | DUT IQ 2 ^{ème} année | 12h | | 16h |
| Total | | | 24h | 140h | |

TABLE 1: Récapitulatif des enseignements effectués

1.2 Détails des enseignements

Conception orientée objets L'objectif de ces travaux pratiques (TP) a été d'initier les étudiants de première année DUT informatique à la modélisation orientée objet. les principes de la programmation orientée-objet tels que : l'encapsulation, l'héritage et le polymorphisme ont été abordés.

De façon plus spécifique, ce cours a permis aux étudiants de :

- Maîtriser une suite de modélisation UML comme *Visual Paradigm* ;
- Découvrir le développement de manière collaborative de logiciels en utilisant des systèmes de version contrôle et agile ;
- Maîtriser les concepts de la programmation orientée objet en utilisant le langage Java.

Architecture des systèmes L'objectif de ces travaux pratiques a été de permettre aux étudiants de première année DUT informatique, ayant déjà des connaissances de base en programmation C, de maîtriser la programmation bas niveau, afin de concevoir et commander des systèmes d'entrée-sortie dans un environnement de microcontrôleur simulé.

Programmation mobile Les 12h de travaux dirigés (TD), et les 16h de TPs ont permis aux étudiants de deuxième année DUT informatique de développer des applications mobiles sous Android afin d'apprendre à créer une application native pour téléphone ou tablette.

Programmation web orienté client Les 12h de TDs et les 32h de TPs ont été dispensés aux étudiants de deuxième année DUT informatique pour acquérir des connaissances nécessaires au développement d' :

- un jeu web programmé en javascript ;
- une interface web pour permettre la communication avec un logiciel d'ERP développé lors d'un autre cours.

1.3 Supervision de projet

Durant ma période postdoctorale, j'ai eu l'occasion de superviser des étudiants de master dans le laboratoire et de conseiller les doctorants dans les domaines proches de ma recherche.

1.4 Membre actif du project SALEIE

Durant ma période postdoctorale, j'ai l'occasion de travailler dans le cadre du projet Européen H2020 SALEIE (Strategic Alignment of Electrical and Information Engineering in European Higher Education Institutions). Le but du projet a été de confectionner une révision des programmes d'enseignement supérieur européen dans le domaine de l'information et de la communication (TIC). Un résumé concernant ces réflexions a été publié dans [1].

2 Activités de recherches

2.1 Doctorat

- Titre : **Segmentation d'objets déformables en imagerie ultrasonore**
- Institutions : Université de Bourgogne au laboratoire Le2i (Laboratoire d'Electronique, Informatique et Image) / Universitat de Girona à Institut VICOROB
- Période : Octobre 2009 à Décembre 2013
- Soutenue le : 4 Décembre 2013
- Mention : Très Honorable
- Directeur de thèse : **Fabrice Meriaudeau**, Professeur à l'Université de Bourgogne
- Codirecteur de thèse : **Joan Martí**, Professeur à l'Universitat de Girona
- Jury de thèse :

| | | | | |
|----------------------|----------------------|-----------------------------------|--------------------|--------|
| Denis Friboulet | Professeur | INSA Lyon (Creatis) | Président du jury | CNU 61 |
| Robert Martí | Maitre de conférence | Universitat de Girona (Vicorob) | Co-directeur | - |
| Fabrice Meriaudeau | Professeur | Université de Bourgogne (Le2i) | Directeur de thèse | CNU 61 |
| Francesco Tortorella | Professeur | Università degli Studi di Cassino | Directeur de thèse | - |

2.1.1 Résumé de thèse

Le cancer du sein est la cause principale de mortalité par cancer chez les femmes. Bien que la Mammographie Numérique (MN) reste la référence pour les méthodes d'examen existantes, l'imagerie ultrasonore a prouvé son efficacité en tant que modalité complémentaire, et on estime qu'elle pourrait éviter 65 à 85% des biopsies prescrites. Cependant, les images ultrasonores sont difficilement interprétables, c'est pour cela que la communauté médicale a mis au point un lexique commun réduisant les incohérences entre radiologues. Une telle pratique est énormément couteuse en temps.

Les systèmes de diagnostic assisté par ordinateur (DAO) ont été développés afin d'aider les radiologues dans la prise décision concernant les lésions détectées. Cependant, ces systèmes ne prennent pas en compte le lexique développé par ces derniers, ce qui rend leurs utilisations compliquées.

Mes travaux de thèse ont eu pour but de concevoir un DAO compatible avec le lexique mise en place par les médecins. Une analyse du processus de segmentation est effectuée et une nouvelle méthode automatique de segmentation sur des images ultrasons (US) est proposée.

2.1.2 Les contributions

Base de données publique d'images ultrasonores du sein Créer une base de données publique d'images d'échographie du sein. De plus, les lésions dans ces images ont été segmentées et annotées par des radiologues.

Proposition de deux nouvelles méthodologies En plus d'une révision approfondie des méthodologies existantes [2], j'ai proposé deux méthodes pour la segmentation de lésions dans les images d'échographie :

- (i) l'une est basée sur la propagation d'un front d'onde en utilisant un modèle Gaussien [3, 4] ;
- (ii) l'autre met en œuvre une méthode de minimisation basée sur des coupes de graphe combinée avec une représentation en superpixels donnant lieu aux publications suivantes [7, 6, 5].

2.2 Post-doctorat

J'ai effectué un post-doc dans le groupe d'imagerie médicale du laboratoire Le2i à l'Université de Bourgogne. En collaboration avec les autres chercheurs de l'équipe, j'ai pris part aux différents projets tels que la détection des mélanomes, la segmentation de cancer de la prostate, la segmentation de cancer du sein ainsi que les maladies rétiennes. Comme mentionné précédemment, l'une de mes activités a été de conseiller dans leurs recherches les doctorants du groupe.

Classification automatique d'Eudème Maculaire Diabétique (EMD). Cette recherche est liée à un Projet Hubert Curien (PHC) permettant une collaboration avec le Singapore Research Eye Institute (SERI) afin de créer des outils pour la détection automatique de maladies rétiennes. Durant ces deux années, nous avons proposé deux méthodes pour la détection de patients atteints d'EMD dans les images de Tomographie en Cohérence Optique (TCO) : (i) une méthode supervisée utilisant des descripteurs de texture un modèle de sacs de mots visuels [8, 9] et (ii) une méthode semi-supervisée basée sur un modèle de mélange de gaussiennes [10]. De plus, nous avons également proposé un état-de-l'art et un banc de tests pour les méthodes existantes publiés dans [11].

Développement de systèmes DAO pour des applications en imagerie médicale J'ai également contribué au développement des systèmes de DAO créés dans les différentes applications en imagerie médicale au sein de notre laboratoire : détection automatique d'EMD et de mélanome et segmentation automatique de lésions de cancer de la prostate et du sein. J'ai plus particulièrement travaillé sur la manière de généraliser les méthodes d'apprentissage automatique pour pouvoir les appliquer dans ces différentes applications donnant lieu aux publications suivantes [12, 13, 14, 15, 16, 8, 17, 6, 9, 18, 14].

En outre de ma recherche en imagerie médicale, j'ai également pris part à des projets industriels contractés par mon laboratoire. Un de ces projets a également permis une publication scientifique de la preuve de concept en utilisant l'imagerie non-conventionnelle pour trouver l'orientation de fibres [20, 19].

2.3 Perspective de recherche

J'ai eu l'occasion d'exceller dans le domaine de l'apprentissage statistique et automatique, de la reconnaissance de forme, du traitement d'images notamment en imagerie médicale, mais également en imagerie non conventionnelle.

3 Autres activités

3.1 Organisation d'évènements scientifiques

J'ai été co-organisateur de la deuxième édition du Doctoral Day 2015, organisée au Creusot.

J'ai également participé à l'organisation à la semaine d'intégration, le Vibot Day ainsi que la remise des diplômes du Master Erasmus Mundus Vibot.

J'ai aussi créé un groupe de travail dans le laboratoire Le2i qui se réunit régulièrement pour la lecture de travaux scientifiques.

3.2 Culture scientifique

J'ai travaillé sur divers projets de recherche pédagogique tels que SALARIÉ et le projet Européen Erasmus+ Playful Coding, donnant lieu aux publications suivantes [21, 22].

3.3 Relecture d'articles scientifiques

J'ai effectué des relectures pour des revues scientifiques et de conférences internationales.

4 Publications

Toutes mes revues publiées sont toutes référencées JCR.

Revues internationales

- [9] Guillaume LEMAÎTRE, Mojdeh RASTGOO, Joan MASSICH, Carol Y CHEUNG, Tien Y WONG, Ecosse LAMOUREUX, Dan MILEA, Fabrice MERIAUDEAU et Désiré SIDIBÉ. "Classification of SD-OCT Volumes using Local Binary Patterns: Experimental Validation for DME Detection". In : *Journal of Ophthalmology* 2016 (2016).
- [10] Désiré SIDIBÉ, Shrinivasan SANKAR, Guillaume LEMAÎTRE, Mojdeh RASTGOO, Joan MASSICH, Carol Y CHEUNG, Gavin SW TAN, Dan MILEA, Ecosse LAMOUREUX, Tien Y WONG et al. "An anomaly detection approach for the identification of DME patients using spectral domain optical coherence tomography images". In : *Computer Methods and Programs in Biomedicine* 139 (2017), p. 109–117.
- [19] Mohamed BELKACEMI, Christophe STOLZ, Alex andre MATHIEU, Guillaume LEMAÎTRE, Joan MASSICH et Olivier AUBRETON. "Nondestructive testing based on scanning-from-heating approach: application to nonthrough defect detection and fiber orientation assessment". In : *Journal of Electronic Imaging* 24.6 (2015), p. 061112–061112.

Conférences internationales

- [1] Jana LIGUSOVA, Nina BENCHEVA, Jean-Marc THIRIET, Gert JERVAN et Massich JOAN. "Reflections about the integration of global challenges into higher education future programs: application in the field of ICT security". In : *Proceedings of ITHERET 2014. 13th International Conference on Information Technology Based Higher Education and Training*. Sept. 2014.
- [3] Joan MASSICH, Fabrice MERIAUDEAU, Elsa PÉREZ, Robert MARTÍ, Arnau OLIVER et Joan MARTÍ. "Lesion Segmentation in Breast Sonography". English. In : *Digital Mammography*. Sous la dir. de Joan MARTÀ, Arnau OLIVER, Jordi FREIXENET et Robert MARTÀ. T. 6136. Lecture Notes in Computer Science. Springer Berlin Heidelberg, juin 2010, p. 39–45. ISBN : 978-3-642-13665-8.
- [4] Joan MASSICH, Fabrice MERIAUDEAU, Elsa PÉREZ, Robert MARTÍ, Arnau OLIVER et Joan MARTÍ. "Seed selection criteria for breast lesion segmentation in ultra-sound images". In : *Proc. MICCAI 2011 Workshop on Breast Image Analysis*. Sous la dir. de Christine TANNER, Julia SCHNABEL, Nico KARSSENMEIJER, Mads NIELSEN, Maryellen GIGER et Dawid HAWKES. Department of computer siccience (DIKU), University of Copenhagen, sept. 2011, p. 57–64. ISBN : 978-87-981270-9-3.
- [5] Joan MASSICH, Guillaume LEMAÎTRE, Joan MARTÍ et Fabrice MERIAUDEAU. "An optimization approach to segment breast lesions in ultra-sound images using clinically validated visual cues". English. In : *Proc. MICCAI 2015 Workshop on Breast Image Analysis (BIA)*. Munich, Germany, oct. 2015.
- [6] Joan MASSICH, Fabrice MERIAUDEAU, Melcior SENTÍS, Sergi GANAU, Elsa PÉREZ, Domènec PUIG, Robert MARTÍ, Arnau OLIVER et Joan MARTÍ. "SIFT Texture Description for Understanding Breast Ultrasound Images". English. In : *Breast Imaging*. Sous la dir. d'Hiroshi FUJITA, Takeshi HARA et Chisako MURAMATSU. T. 8539. Lecture Notes in Computer Science. Springer International Publishing, 2014, p. 681–688. ISBN : 978-3-319-07886-1.
- [7] Joan MASSICH, Fabrice MERIAUDEAU, Melció SANTÍS, Sergi GANAU, Elsa PÉREZ, Robert MARTÍ, Arnau OLIVER et Joan MARTÍ. "Automatic seed placement for breast lesion segmentation on US images". English. In : *Breast Imaging*. Sous la dir. d'AndrewD.A. MAIDMENT, PredragR. BAKIC et Sara GAVENONIS. T. 7361. Lecture Notes in Computer Science. Springer Berlin Heidelberg, juil. 2012, p. 308–315. ISBN : 978-3-642-31270-0.
- [8] Guillaume LEMAÎTRE, Mojdeh RASTGOO, Joan MASSICH, Desire SIDIBE et Fabrice MERIAUDEAU. "Classification of SD-OCT volumes with LBP: Application to DME detection". English. In : *Proc. MICCAI 2015 Workshop on Ophthalmic Medical Image Analysis (OMIA)*. Munich, Germany, oct. 2015.
- [11] Joan MASSICH, Mojdeh RASTGOO, Guillaume LEMAÎTRE, Carol CHEUNG, Tien Y WONG, Desire SIDIBE et Fabrice MERIAUDEAU. "Classifying DME vs normal SD-OCT volumes: A review". English. In : *23rd International Conference on Pattern Recognition (ICPR)*. Cancun: Mexico, déc. 2016.

- [12] Khaled ALSAIH, Guillaume LEMAÎTRE, Joan MASSICH, Mojdeh RASTGOO, Désiré SIDIBÉ, Tien Y WONG, Ecosse LAMOUREUX, Dan MILEA, Carol Y CHEUNG et Fabrice MÉRIAUDEAU. "Classification of SD-OCT volumes with multi pyramids, LBP and HOG descriptors: application to DME detections". In : *38th IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016.
- [13] Mojdeh RASTGOO, Guillaume LEMAÎTRE, Joan MASSICH, Oliver MOREL, Frank MARZANI, Rafael GARCIA et Fabrice MÉRIAUDEAU. "Tackling the Problem of Data Imbalancing for Melanoma Classification". English. In : *3rd International conference on BIOIMAGING*. Rome, Italy, fév. 2016.
- [14] Guillaume LEMAÎTRE, Mojdeh RASTGOO, Joan MASSICH, Joan C. VILANOVA, Paul M. WALKER, Jordi FREIXENET, Anke MEYER-BAESE, Robert MARTÍ et Fabrice MÉRIAUDEAU. "Normalization of T2W-MRI Prostate Images using Rician a priori". English. In : *SPIE Medical Imaging 2016*. S, fév. 2016, p. 978529–978529.
- [15] Mojdeh RASTGOO, Guillaume LEMAÎTRE, Oliver MOREL, Joan MASSICH, Frank MARZANI, Rafael GARCIA et Desire SIDIBE. "Classification of melanoma lesions using sparse coded features and random forests". English. In : *SPIE Medical Imaging 2016*. S, fév. 2016.
- [16] Anke MEYER-BAESE, Joan MASSICH, Guillaume LEMAÎTRE et Mojdeh RASTGOO. "Real-Time Optical Flow with Theoretically Justified Warping Applied to Medical Imaging". English. In : *Proc. MICCAI 2015 Workshop on Ophthalmic Medical Image Analysis (OMIA)*. Munich, Germany, oct. 2015.
- [17] Guillaume LEMAÎTRE, Joan MASSICH, Robert MARTÍ, Freixenet JORDI, J.C. VILANOVA, P.M. WALKER, Desire SIDIBE et Fabrice MÉRIAUDEAU. "A Boosting Approach for Prostate Cancer Detection using Multi-parametric MRI". English. In : *Proc. International Conference on Quality Ciontrol and Artificial Vision (QCAV)*. Le Creusot, France, juin 2015.
- [18] Mojdeh RASTGO, Guillaume LEMAÎTRE, Olivier MOREL, Joan MASSICH, Rafael GARCIA, Fabrice MERIAUDEAU, Franck MARZANI et Désiré SIDIBÉ. "Classification of melanoma lesions using sparse coded features and random forests". In : *SPIE Medical Imaging*. International Society for Optics et Photonics. 2016, p. 97850C–97850C.
- [20] M BELKACEMI, J MASSICH, G LEMAÎTRE, C STOLZ, V DAVAL, G POT, O AUBRETON, R COLLET et F MERIAUDEAU. "Wood fiber orientation assessment based on punctual laser beam excitation". English. In : *13rd Quantitative Infrared Thermography Conference (QIRT)*. Gdansk, Poland, juin 2016.
- [21] Miquel VILLANUEVA, Xevi CUFÍ, Andrés ELFAKDI, Joan MASSICH et Rafael GARCIA. "Attracting talent to increase interest for engineering among secondary school students". In : *Global Engineering Education Conference (EDUCON), 2011 IEEE*. IEEE. 2011, p. 347–353.
- [22] Xevi CUFÍ, Miquel VILLANUEVA, Andrés ELFAKDI, Joan MASSICH et Rafael GARCIA. "Team-based Building of a Remotely Operated Robot as a Method to Increase the Interest for Engineering among Secondary School Students". In : *Proceedings of EDULEARN 2012. 4th International Conference on Education and New Learning Technologies*. Juil. 2012.
- [23] Guillaume LEMAÎTRE, A. BIKFALVI, J. LLACH, Joan MASSICH et F. JULIAN. "Business Model Design for University Technology Valorisation". English. In : *Proc. International Technology, Education and Development Conference (INTED)*. Madrid, Spain.
- [24] Joan MASSICH, Guillaume LEMAÎTRE, Fabrice MERIAUDEAU et Joan MARTÍ. "Breast Ultra-Sound Image Segmentation: an Optimization approach based on super-pixels and high-level descriptors". English. In : *Proc. International Conference on Quality Ciontrol and Artificial Vision (QCAV)*. Le Creusot, France, juin 2015.
- [29] Emili HERNÀNDEZ, Pere RIDAO, Marc CARRERAS, David RIBAS, Narcís PALOMERAS, Andrés ELFAKDI, François CHUNG, Xavier RIBAS, Guillermo GARCÍA DE MARINA, Natalia HURTÓS, Joan MASSICH, Antonio ALMOHAYA et Josep VILA. "ICTINEU AUV, un Robot per a Competir". Catalan. In : *Artificial Intelligence Research and Development, Proceedings of the 9th International Conference of the ACIA, CCIA 2006*. Sous la dir. de Monique POLIT, Thierry TALBERT, Beatriz LÓPEZ et Joaquím MELÉNDEZ. T. 146. Frontiers in Artificial Intelligence and Applications. IOS Press, oct. 2006. ISBN : 978-1-58603-663-8.

Chapitre de livres

- [25] Joan MARTÍ, GUBERN-MÉRIDA, Joan MASSICH, Arnau OLIVER, Joan C. VILANOVA, Josep COMET, Elsa PÉREZ, Arzoz M et Robert MARTÍ. "Ultrasound Image Analysis. Methods and Applications." In : *Recent advances in biomedical signal processing*. Sous la dir. de Juan Manuel GÓRRIZ, Elmar W LANG et Javier RAMÍREZ. Bentham Science Publishers, 2011, p. 216–230.

Divers

- [26] Gabriel FALCAO, Natalia HURTOS et Joan MASSICH. *Plane-based calibration of a projector-camera system*. Rapp. scient. Le Creusot, France : Shape Recognition Statistics course, déc. 2008.
- [27] Gabriel FALCAO, Natalia HURTOS, Joan MASSICH et David FOFI. *Projector-camera calibration toolbox*. Logiciel. Le Creusot, France, fév. 2009. URL : <http://code.google.com/p/procamcalib>.
- [28] David RIBAS, Narcís PALOMERAS, Xavier RIBAS, Guillermo GARCÍA DE MARINA, Emili HERNÁNDEZ, Fran ois CHUNG, Natalia HURT S, Joan MASSICH, Antonio ALMOHAYA, Josep VILA et Andr s EL-FAKDI. *ICTINEU AUV Takes the Challenge*. Rapp. scient. Girona, Catalonia, 2006.

5 Annexes

Les documents suivants sont joints à ce dossier en annexe :

- Attestation et recommandation de Cédric Demonceaux, Professeur, responsable du site du Creusot - Le2i.
- Attestation et recommandation de Sylvain Rampacek, Maître de Conférence, chef du département informatique de l'IUT de Dijon.
- Recommandation de Joan Martí, Professeur à l'Universitat de Girona.
- Pré-rapports confidentiels de thèse. Ces rapports sont confidentiels du fait des accords de co-tutelle entre l'Université de Bourgogne et l'Universitat de Girona, appliquant les règles de cette dernière.
- Attestion de réussite au diplome de thèse.
- 2 Publications en tant que premier auteur.



Le2i UMR CNRS 6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté

Le Creusot, le 25 mars 2016

Objet : Lettre de recommandation pour le recrutement de Joan Massich sur un poste de maître de conférences.

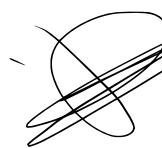
Joan Massich était sur un contrat de post-doctorat sur le site du Creusot du laboratoire Le2i pendant deux ans. A ce titre, il a mené des recherches en traitement des images médicales sur lesquelles il applique des méthodes de *machine learning* pour la détection et la reconnaissance de maladies. Depuis mars 2016, il est attaché temporaire et de recherche à l'IUT de Dijon-Auxerre département Informatique. Malgré la lourde charge d'enseignement qui lui incombe et les déplacements importants entre Dijon et Le Creusot, Joan continue à mener ses recherches de manière très intensive.

Dès son arrivée sur notre site, Joan a créé un groupe de travail qui se réunit toutes les deux semaines. Ce groupe de travail consiste à analyser un article sur lequel un des collègues est en train de travailler pour l'aider l'appréhender. Cette idée de créer un tel groupe de travail montre que Joan a un esprit d'équipe très prononcé et qui me semble indispensable dans une équipe de recherche. Ce groupe de travail a d'ailleurs souvent permis des avancées significatives pour les recherches menées par les collègues et a amélioré l'interaction entre les chercheurs de l'équipe creusotine.

En tant que collègue, Joan est une personne à l'écoute, toujours prête à aider ses collègues pour des tâches tant scientifiques qu'administratives. Je suis sûr que Joan Massich sera un excellent maître de conférences, capable de mener des recherches de qualité sans négliger sa mission d'enseignant. C'est pourquoi, je me permets d'appuyer très favorablement la candidature de Joan Massich pour un poste de Maître de Conférences.

Je reste à votre disposition si vous souhaitez de plus amples informations à son sujet.

Cédric Demonceaux
Professeur des Universités
Responsable du site du Creusot – Le2i
Responsable du pôle 6,Vision pour la Robotique
Le2i UMR CNRS 6306



Laboratoire Le2i
Unité Mixte de Recherche CNRS 6306 / Arts et Métiers / Univ. Bourgogne Franche-Comté
IUT, 12 rue de la Fonderie, 71200 LE CREUSOT
Tél. : 03 85 73 10 90 - Fax. : 03 85 73 10 97
<http://le2i.cnrs.fr>

Sylvain Rampacek
Chef du département informatique de l'IUT de Dijon-Auxerre
Maître de Conférences en Informatique

Dijon, le 16 décembre 2016

A qui de droit,

Object Recommandation enseignement – M. Massich Joan

Madame, Monsieur,

Durant l'année 2015/2016, Massich Joan a assuré des enseignements, en tant que ATER, dans les modules d'Architecture des systèmes (20h TP), de Conception Orientée Objet (72h TP), de Programmation Web Orienté Client (12h TD et 32h TP) et de Programmation mobile (12h TD et 16h TP), au niveau du DUT Informatique de l'IUT de Dijon-Auxerre.

Au cours de tous ses enseignements, Massich Joan a fait preuve d'autonomie. Il a montré ses qualités en préparent ces enseignements et est une personne apprécié par les étudiants.

Pour toutes ces raisons, je pense que Massich Joan satisfait les principaux critères, en enseignement, pour son inscription sur la liste de qualification aux fonctions de Maître de Conférences.

Sylvain Rampacek
Chef du Département Informatique



April 11th, 2016
Computer Vision and Robotics Institute
University of Girona
Campus de Montilivi s/n
17071 Girona

To whom it may concern,

I am pleased to write this letter to recommend Dr. Joan Massich, currently a researcher at the Université de Bourgogne, as part of his application for a postdoctoral position. Dr. Joan Massich has made a Bachelor Degree and a Master Degree in Computer Science at the University of Girona, where he received excellent grades. Once he finished his studies he joined the Erasmus Mundus Master in Computer Vision & Robotics (VIBOT) held by our University of Girona (Spain), and the Heriot-Watt University (UK) and Université de Bourgogne (France). During this period, he was holding a Scholarship by the European Commission.

After the VIBOT Master, Dr. Joan Massich enrolled our Computer Vision and Robotics group, where he achieved the PhD degree under my supervision, with the work entitled "*Deformable object segmentation in ultrasound images*". I would like to remark that his performance has been at a top level, carrying out his tasks with accuracy and dedication. It has been really an easy task to work with him, provided he has always shown receptive and eager to use the advices he receives from other colleagues. Observing him during this period, I have to say that he has a great capacity for doing research and he has innate leadership skills. He is a very active person, who always feels comfortable with new challenges. His work has been accurate and with an outstanding quality level, always achieved in an efficient way.

Dr. Joan Massich has always shown his interest in doing research and continuing with an academic career. I honestly believe that this opportunity could be very useful for him. I have no doubts that he will benefit from this opportunity working at a top level.

Please, do not hesitate to contact me if you need more information or would like me to answer additional questions.

Yours Sincerely,



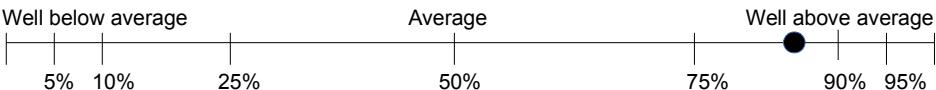
Dr. Joan Martí
E-mail: joanm@eia.udg.edu
www: <http://eia.udg.edu/~joanm>

DOCTORAL THESIS REVIEW REPORT

| Information on the doctoral thesis | |
|------------------------------------|--|
| Title | Deformable object segmentation in Ultra-Sound images |
| Submitted by | Joan Massich |

Global assessment: 85 %

Well below average Average Well above average



(Indicate the relative position of the thesis to other theses in the same area.)

Do you think the thesis can be defended?

Without changes
 With minor changes
 With major changes
 Not in its current version

Do you think the thesis deserves the distinction “cum laude”? No
(As a reference, if you consider that the Thesis is above average you should answer Yes to this question)

REASONED REVIEW (the reviewer has to justify the global assessment and the proposed changes in the document in terms of the technical aspects and results (theoretical framework, relevance, objectives, methodology, discussion, conclusions, bibliography, worth publishing) and in terms of the formal aspects (presentation, how well it is written, spelling...)).

Summary and recommendation

Joan Massich's thesis addresses the highly topical and challenging area of automatic analysis of *Breast Ultra-Sound* (BUS) images. This is particularly interesting and important when it is considered in the framework of massive screening where a system for automatic analysis could improve the detection of a breast cancer at a very early stage, when it is still possible to successfully attack the disease with a suitable therapy. In particular, BUS imaging has proven to be a successful adjunct

image modality for breast cancer screening with respect to other reference technique such as the *Digital Mammography* (DM), specially in cases where the DM is not sufficiently reliable (e.g. very dense breast).

The thesis gives a valuable contribution to this field. In the first part (chapters 1, 2) the author introduces the imaging modalities used in breast screening placing special emphasis in US screening of the breast and provides a state-of-the-art in segmentation of breast lesions in US data. In particular, he points out some problems that make difficult a fair comparison between different methods such as the lack of a common dataset and of common assessment measures.

In the second part of the thesis (chapters 3, 4), the author firstly reviews the segmentation method based on *Gaussian Constrain Segmentation* presented in [J.Massich, F.Meriaudeau, E.Pérez, R.Martí, A.Oliver, and J.Martí. *Lesion segmentation in breast sonography*. International Workshop on Digital Mammography, LNCS 6136, pp 39-45. Girona, Spain. June 2010] and in [J.Massich, F.Meriaudeau, M.Sentís, S.Ganau, E.Pérez, R.Martí, A.Oliver, and J.Martí. *Automatic seed placement for breast lesion segmentation on US images*. International Workshop on Digital Mammography, LNCS 7361, pp 308-315. Philadelphia, Pennsylvania. July 2012]. Subsequently, he introduces a novel segmentation scheme based on an optimization framework. In particular, for each image a first-level segmentation is performed in order to decompose the image in groups of contiguous pixels sharing some characteristic (*superpixels*). Each superpixel is then characterized by means of a set of features extracted from the images and fed to a classifier trained to recognize the superpixels belonging to a lesion. The final labeling of the superpixels takes into account both the output of the classifier and the relationship among near superpixels so as to ensure spatial coherence in the final segmentation; this is made so as to minimize a cost related to the configuration assumed by the labels in the final segmentation.

For each of the steps described above, the author delineates some possible solutions and discusses pros and cons for each of them, thus providing a flexible framework in which each subtask can be designed individually. The proposed approach has been tested on a dataset of real BUS images: several possible configurations of the general method have been considered and their results have been compared.

The obtained results demonstrate the potential of the proposed approach and encourage to take this work further.

Presentation

The thesis is well presented, and interesting to read.

Contents

The thesis represents a large body of work and is innovative in several areas.

LIST OF QUESTIONS AND COMMENTS

General comments

- Sect. 3.2.2: some details about how the PDFs are estimated from training data would be useful. Is the estimate parametric or non parametric (e.g. Parzen) ?
- Sect. 3.3.2: it is said that the speckle noise is not removed because it could be profitable for the classification process. Could the author motivate such statement ? There is some reference in literature or some experimental evidence ?
- Sect. 3.3.4: the features describing the brightness of the regions apparently are not normalized (see eqq. 3.5 and 3.6) while, if figs. 3.12 and 3.14 are considered, they appear to be normalized between 0 and 1. Could the author clarify this point ?
- Sect. 3.3.4 (Describing the overall appearance of the region): the description of the features is not very clear and could be improved. What is a MAD brightness model and how is it calculated ? Which kind of models are defined ? Are they tissue models (muscle, pleura, fat, ...) or echo pattern models ?
- Sect. 3.3.4 (Describing the texture appearance of the regions): To generate the visual dictionary the k-means clustering is used. How is chosen the value for k ?
- Sect. 3.3.4 (Describing the location of the region): How is managed the atlas option ? Is it a feature added to the other features ? Or is it combined with the output of the classifier ?

- Sect. 3.3.5: In this section several criteria to select the most adequate classification technique are described. How such criteria have lead to the choice of SVM RBF and how is the parameter σ chosen ? Did the author consider other classification architectures ?
- Sect. 3.5 (Quantitative results): the cross-validation is carried out as a multiple randomized sampling of the pool of superpixels. Which is the size of the training set and of the test set in each round of the cross-validation ? How many rounds are performed ?

Minor comments and typos

- p. 10: “Position Emission Tomography” should be “Positron Emission Tomography”
- p. 28: “projectes” should be “projects”
- p. 47: “weather” should be “wether”
- p. 49 (caption fig. .5): maybe “quantified” should be “quantized”
- p. 50: “unpropper detection” should be “improper detection”
- p. 52: the last term of eq. 2.1 should be $\frac{|A \wedge M|}{|A \vee M|}$
- p. 53 (eqs. 2.4 and 2.5): the term $|A| \wedge |M|$ should be $|A \wedge M|$
- p. 54: “Precision” should be “Precision”
- p. 54: “the inverse of the TPR” should be “the complement of the TPR”
- p. 57: “Multipel grader” should be “Multiple grader”
- p. 72 (fig. 3.6): it is not clear which is the method from that the images in the figure have been extracted,
- p.73: “while the later” should be “while the latter”
- p. 74: “part of the designing process” should be “part of the design process”
- p. 82: the eqs. 3.5, 3.6 are totally obscure: the notation is not clear and there are several symbols used but not introduced before (e.g., $\mu(\cdot)$ and $Md(\cdot)$).
- p. 83: “Figures 3.13 and 3.14 replicate the study for a different image example ...” should be “Figures 3.14 and 3.15 replicate the study for a different image example ...”
- p. 102: “biological behabiours ...” should be “biological behaviours ...”
- p. 111: in *Superpixel multi-resolution feature description* there are two median descriptors while one of them should be a mean descriptor.
- p. 117: “more profound” should be “deeper”
- p. 118: “Figures 3.38 to 3.40 show few quantitative results ...” should be “Figures 3.38 to 3.40 show few qualitative results ...”.
- pp. 119-120 (figs 3.36 and 3.37): in the graph the FN ring is not present.
- Bibliography: some references are incomplete



DOCTORAL THESIS REVIEW REPORT

| Information on the doctoral thesis | |
|--|--|
| Title Deformable object segmentation in Ultra-Sound images | |
| Submitted by Mr. Joan Massich | |

| |
|---|
| Global assessment: 70% |
| A horizontal scale from 5% to 95% with markers at 5%, 10%, 25%, 50%, 75%, 90%, and 95%. The scale is divided into three main segments: "Well below average" (0-25%), "Average" (25-75%), and "Well above average" (75-95%). |
| (Indicate the relative position of the thesis to other theses in the same area.) |

| | |
|---|--|
| Do you think the thesis can be defended? | <input type="checkbox"/> Without changes |
| | <input checked="" type="checkbox"/> With minor changes |
| | <input type="checkbox"/> With major changes |
| | <input type="checkbox"/> Not in its current version |

| | | |
|---|---|-----------------------------|
| Do you think the thesis deserves the distinction “cum laude”? | <input checked="" type="checkbox"/> Yes | <input type="checkbox"/> No |
| (As a reference, if you consider that the Thesis is above average you should answer Yes to this question) | | |

REASONED REVIEW (the reviewer has to justify the global assessment and the proposed changes in the document in terms of the technical aspects and results (theoretical framework, relevance, objectives, methodology, discussion, conclusions, bibliography, worth publishing) and in terms of the formal aspects (presentation, how well it is written, spelling...)).

Thesis report for « Deformable object segmentation in Ultra-Sound images »
by Joan MASSICH

The applicative context of this thesis concerns segmentation of breast cancer lesions in ultrasound echography. This is a hard and state-of-art problem in the field of medical image analysis, due to the difficulties linked to ultrasonic image formation (inducing low contrast, speckle, shadowing, etc.). In terms of methodology, the objective of Mr. Joan Massich is the development of a fully automated segmentation method based on machine learning and optimization techniques.

The manuscript presented by Mr. Joan Massich is composed of 4 chapters, i.e. an introduction giving the applicative context, a review of lesion segmentation methods in ultrasound images, the description of the main contribution and a conclusion. The manuscript is clear and well-written. I provide a chapter-by-chapter analysis below.

Chapter 1 provides a detailed introduction to the aim of the thesis work. Starting with the description of the role of different imaging modalities for the assessment of breast cancer, it then focuses on ultrasound imaging and provides an extensive description of the specific role of this modality in breast cancer detection and staging. The chapter then introduces the notion of CAD in breast lesion detection and the prominent role of segmentation in this context. This chapter is clear, uses a rich iconography and constitutes a very nice introduction to the problematic of breast cancer imaging.

Chapter 2 is background review of lesion segmentation in ultrasound imaging. The literature covered by this chapter is extensive and the author first proposes to structure it through well-defined categories, i.e. the degree of interaction involved, the core methodology employed (i.e. active contours, machine learning, etc.) and the image features used. One important aspect of segmentation is the quantitative evaluation of the results, and the author details quite extensively the assessment criteria used in the literature to compare a given segmentation result with a ground truth. The chapter ends with a discussion analyzing previous studies dealing with breast lesion segmentation in terms of methodology and quantitative performance, using the Area overlap (AOV) as main criteria. While this chapter is overall extensive and adequate, it is a bit regrettable that the work done by the author (reference [61]) is not clearly in the next chapter. Moreover, it would have been interesting to add at this level a section giving some overall conclusion about the review and explaining the methodological directions chosen with regards to this review.

Chapter 3 represents the core of the thesis, since it describes the contributions developed by Mr. Joan Massich for the segmentation of breast lesion.

This chapter starts by presenting a first method called "Gaussian constrained segmentation" (GCS). The first step of GCS consists in generating a seed region, obtained by computing the posterior probability of the pixels of belonging to a lesion and thresholding the so-obtained probability map. This posterior probability is obtained from prior probabilities of intensity, texture and position of lesion estimated from a set of training, annotated images. A region growing is applied to the seed and a 2D Gaussian is then fitted on the obtained region. As a second step, an image synthetizing the data characteristics (homogeneity, texture, etc.), Ψ , is generated and multiplied by this 2D Gaussian. The final segmentation is then obtained by thresholding the resulting image, the threshold being optimized to minimize the variance of the lesion and background. Qualitative results provided by GCS are provided and the reader is referred to chapter 2 for a quantitative evaluation, which shows that the AOV associated to GCS 64%, which represents a performance that could be improved. While the overall approach underlying GCS approach is clear, its description suffers from a lack of details, e.g. what criteria is used to stop the region growing, what is the annotated training set, what are the texture features used to compute the prior probability, what are the details of the processing steps (median filtering, morphological operators) used to generate Ψ ? Moreover, it would have been clearer to briefly recall the obtained quantitative results and above all explain from them the need to develop another technique.

The second part of chapter 3 is devoted to the description and evaluation of a second approach. The segmentation task is tackled in this part as a labeling problem, solved using optimization techniques. In order to decrease the complexity, the sites considered for labeling are superpixels, built using the Quick-Shift or the Global Probability Boundary techniques. A SVM classifier is used to generate the labeling cost of a given superpixel and the optimization step then consists in minimizing the labeling cost plus an original smoothing term, enforcing similar labeling of spatially neighboring superpixels. This non convex optimization problem is solved using graph cut. The whole procedure can thus be seen as an a posteriori regularization of the SVM output: the expected solution should indeed be close to the SVM classification (data term) while constraining spatially neighboring sites to have similar labels. While the overall structure of the approach is clear, the construction of the data term from the SVM output should have been more clearly detailed: what is the width of the RBF used? How is the cost formulated from the hinge cost function, etc.? In the same way, the details of the construction of the smoothing term based on Markov Random field techniques should have been more accurately exposed.

Being based of a machine learning approach the selection of the features used for the classification/cost generation is an important issue. Intensity, texture and localization are used in this work. Consistently with the Stavros criteria widely accepted in the echographic breast imaging community, the author first introduces an original feature mimicking the evaluation of echographic images as anechoic, hypoechoic, etc. This feature is complemented by a more detailed description of the intensity and corresponding to the distance of a given superpixel to the tissue classes, measured by comparing the associated histogram through the Quadratic Chi distance. The texture features are based on a descriptor provided by SIFT technique, i.e. namely on an 8-bins histogram of the local gradients orientation. Each superpixel is then described through a "Bag of features" using a dictionary of 36 SIFT words learned from training data. The

intensity and texture features are embedded in a multiresolution scheme, obtained by computing the related quantities from the site and its neighbors. Resolution can thus be chosen by varying the size of the neighborhood and 3 levels are used in this work. The last feature is based on the probability of a given spatial location to belong to a particular tissue, this posterior probability model being built from the training set. This part of chapter 3 nicely introduces computer vision techniques which are not so common in medical image processing using a rich iconography and clearly illustrates the associated concepts.

The approach is evaluated from a database consisting of 700 images, using a Leave-One-Out strategy. This evaluation is quite thorough: the grouping of several features is tested through 8 experiments as well as the individual influence of each feature. These experiments show that the approach does not bring a substantial improvement in terms of AOV, whatever the feature configuration (AOV smaller than 61%). Interestingly, the introduction of the proposed regularization term is shown to bring a noticeable reduction of the False Positive rate, at the cost of a slight increase of the False Negative rate. While the work associated to this section is interesting and substantial, it would have been informative to add a discussion comparing in details the obtained results with the GCS and other state-of-the-art methods. In particular some interpretations and perspectives regarding the relatively low obtained AOV as well as the observed decrease of FP would have been of great interest.

Overall, Mr. Joan Massich has shown in this thesis a very good knowledge of the applicative context of echographic breast imaging. He has performed a structured analysis of the existing literature about segmentation of breast lesion. His methodological contributions are twofold. He has first proposed an improvement of the Gaussian constrained segmentation method and has tested this approach of a limited set of 25 images. Mr. Joan Massich has published this aspect of his work in three international conferences proceedings. Mr. Joan Massich has then proposed a computer vision-based approach, relying on the optimization of a cost function including a data term from SVM classification and an original term enforcing the spatial smoothness of labeling. The presentation of this aspect could have been improved by providing more details about the core items of the method. The proposed approach has been thoroughly evaluated on a database of 700 images. While some aspects of the results are clearly improvable (AOV), the approach succeeds in decreasing the false positive rate, which is a common problem in computer aided systems when localizing classification approaches. It is also to be noted that the work

As a conclusion, and based on the above remarks, I thus recommend the oral defense of Mr. Joan Massich for the PhD degree.

UNIVERSITE DE DIJON

ATTESTATION DE REUSSITE AU DIPLOME

La Chargée de scolarité du bureau des doctorants atteste que

le Diplôme de docteur EN INSTRUMENTATION ET INFORMATIQUE DE L'IMAGE
a été décerné à

Monsieur MASSICH VALL JOAN

né le 15 mars 1984 à GERONE (ESPAGNE)

au titre de l'année universitaire 2013/2014 avec la mention Très honorable

Titre des travaux : Segmentation d'objets déformables en imagerie ultrasonore

Date de soutenance : 4 décembre 2013

Etablissement soutenance : ESPAGNE – Université de GERONE

Jury : M. DENIS FRIBOULET, Président du jury, Professeur
M. ROBERT MARTI, Maître de conférence
M. FABRICE MERIAUDEAU, Professeur
M. FRANCESCO TORTORELLA, Professeur

Ecole doctorale : Sciences Physiques pour l'Ingénieur et Microtechniques (SPIM)

Fait à DIJON, le 27 janvier 2014



N° étudiant : 27007702

An optimization approach to segment breast lesions in ultra-sound images using clinically validated visual cues

Joan Massich^{*1}, Guillaume Lemaître^{1,2}, Mojdeh Rastrgooy^{1,2}, Anke Meyer-Baese³, Joan Martí², and Fabrice Mériadeau¹

¹ LE2I-UMR CNRS 6306, Université de Bourgogne, 12 rue de la Fonderie,
71200 Le Creusot, France;

joan.massich@u-bourgogne.fr

² ViCOROB, Universitat de Girona, Campus Montilivi Ed.P4, 17071 Girona, Spain.

³ Department of Scientific Computing, Florida State University,
400 Dirac Science Library Tallahassee, FL 32306, US.

Abstract. As long as breast cancer remains the leading cause of cancer deaths among female population world wide, developing tools to assist radiologists during the diagnosis process is necessary. However, most of the technologies developed in the imaging laboratories are rarely integrated in this assessing process, as they are based on information cues differing from those used by clinicians. In order to grant Computer Aided Diagnosis (CAD) systems with these information cues when performing non-aided diagnosis, better segmentation strategies are needed to automatically produce accurate delineations of the breast structures. This paper proposes a highly modular and flexible framework for segmenting breast tissues and lesions present in Breast Ultra-Sound (BUS) images. This framework relies on an optimization strategy and high-level descriptors designed analogously to the visual cues used by radiologists. The methodology is comprehensively compared to other sixteen published methodologies developed for segmenting lesions in BUS images. The proposed methodology achieves similar results than reported in the state-of-the-art.

Keywords: Breast Ultra-Sound, BI-RADS lexicon, Optimization based Segmentation, Machine-Learning based Segmentation, Graph-Cuts

1 Introduction

Breast cancer is the second most common cancer. In terms of mortality, breast cancer is the fifth most common cause of cancer death. However, it is ranked as

^{*} This work was partially supported by the Regional Council of Burgundy FEDER grant 2013-9201AA0049S02890 and by the Spanish Goverment MEC grant *nb.TIN2012-3171-C02-01*

the leading cause of cancer deaths among females in both western and economically developing countries [4].

Medical imaging contributes to its early detection through screening programs, non-invasive diagnosis, follow-up, and similar procedures. Despite Breast Ultra-Sound (BUS) imaging not being the imaging modality of reference for breast cancer screening [9], Ultra-Sound (US) imaging has more discriminative power when compared with other image modalities to visually differentiate benign from malignant solid lesions [10]. In this manner, US screening is estimated to be able to reduce between 65 ~ 85% of unnecessary biopsies, in favour of a less traumatic short-term screening follow-up using BUS images. As the standard for assessing these BUS images, the American College of Radiology (ACR) proposes the Breast Imaging-Reporting and Data System (BI-RADS) lexicon for BUS images [7]. This US BI-RADS lexicon is a set of standard markers that characterizes the lesions encoding the visual cues found in BUS images and facilitates their analysis. Further details regarding the US BI-RADS lexicon descriptors proposed by the ACR, can be found in Sect. 3, where visual cues of BUS images and breast structures are discussed to define feature descriptors.

The incorporation of US in screening policies and the emergence of clinical standards to assess image like the US BI-RADS lexicon, encourage the development of Computer Aided Diagnosis (CAD) systems using US to be applied to breast cancer diagnosis. However, this clinical assessment using lexicon is not directly applicable to CAD systems. Shortcomings like the location and explicit delineation of the lesions need to be addressed, since those tasks are intrinsically carried out by the radiologists during their visual assessment of the images to infer the lexicon representation of the lesions. Therefore, developing accurate segmentation methodologies for breast lesions and structures is crucial to take advantage of this already validated clinical tools.

2 Description of the segmentation methodology

Optimization methodologies offer a standardized manner to approach segmentation by minimizing an application-driven cost function [2]. Figure 1 illustrates a generic representation of the segmentation strategy, concrete examples of its terms, applied to BUS, can be found in section 3. The overall segmentation can be seen as a three-steps strategy: (1) a mapping of the image into a discrete set of elements \mathcal{S} , (2) the optimization stage which is formulated as a *metric labelling* problem, and (3) a re-mapping the labels obtained from the previous stage to produce the final delineation.

To formulate segmentation as a metric labelling problem, the image is conceived as a discrete set of elements \mathcal{S} that need to be labelled using a label l from the labelling set \mathcal{L} . Let \mathcal{W} be all the possible labelling configurations of the set \mathcal{S} , given \mathcal{L} . Let $U(\cdot)$ be a cost function encoding the goodness of the labelling configuration $\omega \in \mathcal{W}$ based on the appearance of the elements in \mathcal{S} , their inner relation and some designing constraints. Then, the desired segmentation

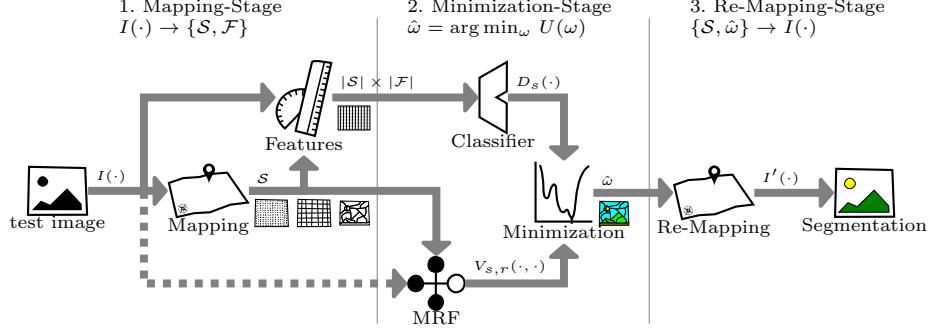


Fig. 1: Conceptual block representation of the segmentation methodology.

$\hat{\omega}$ corresponds to the labelling configuration that minimizes this cost function, as described in Eq. (1).

$$\hat{\omega} = \arg \min_{\omega} U(\omega) \quad (1)$$

This goodness measure $U(\cdot)$ must be defined to take into account the appearance of the target region, its relation with other regions, and other designing constraints. Equation (2) describes this cost function as the combination of two independent costs that need to be simultaneously minimized as a whole.

$$U(\omega) = \sum_{s \in \mathcal{S}} D_s(\omega_s) + \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{N}_s} V_{s,r}(\omega_s, \omega_r) \quad (2)$$

Where, the left-hand side of the expression integrates the so-called *data* term, while the right-hand side integrates the *pairwise* term, which is also referred to as the *smoothing* term. Both terms are shaped by \mathcal{S} and evaluated in the labelling space \mathcal{W} . In our quest to optimize the cost function $U(\cdot)$, it is required to define a representation for the set \mathcal{S} , a data term $D(\cdot)$, a pairwise term $V(\cdot)$, and a proper minimization methodology.

The set \mathcal{S} can be, in general, any discrete set representing the image (i.e. pixels, overlapping or non overlapping windows, super-pixels, etc.).

The *data term* $D(\cdot)$, given a label configuration $\omega \in \mathcal{W}$, penalizes the labelling of a particular image element or site ($\omega_s = l$) based on the data associated to s . In this manner, $D_s(\omega_s = l_{\checkmark}) << D_s(\omega_s = l_{\times})$. Figure 2b illustrates the data cost associated to some arbitrary labelling configurations to clarify the desired effect (or behaviour) of this data term. Designing an obscure heuristic to comply with the desired behaviour of $D(\cdot)$ out of the box, is rather a complicated task. Therefore, an easier and cleaner approach is to design this data term $D(\cdot)$ with the help of Machine Learning (ML) because it provides a systematic process that is flexible enough to encode any desired behaviour based on a training stage. This concept is in fact depicted in the upper row in Fig. 1. For each site $s \in \mathcal{S}$, features

describing s are designed. Then, different optional steps can be applied to this set of features: (i) features normalization, (ii) features selection or (iii) features extraction. Finally, the data term $D(\cdot)$ is encoded based on ML classifiers, the features and a training step. Thus, the data term $D(\cdot)$ can be seen as a distance or goodness measure reflecting the likelihood for s to belong to class l .

The pairwise term $V(\cdot, \cdot)$ represents the cost associated to ω_s taking into account the labels of its neighboring sites, $\omega_r, r \in \mathcal{N}_s$. This term is usually modeled using Markov Random Fields (MRFs) or Conditional Random Fields (CRFs). The typical form of this term, given in Eq. (3), is called homogenization which acts as a regularization factor favouring configurations that have a coherent labelling.

$$V_{s,r}(\omega_s, \omega_r) = \begin{cases} \beta, & \text{if } \omega_s \neq \omega_r \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Figure 2c shows a visual interpretation of this cost. The more fragmented is the segmentation ω , the higher is the overall pairwise term, since every boundary brings a penalization β to the total cost $U(\omega)$. In this manner, the regularization term can be seen as a post-processing or denoising stage as some sites will flip their labelling if the cost of fragmenting the regions is larger than the cost of adopting their neighbour's label.

The minimization strategy is determined by the nature of $U(\cdot)$ and \mathcal{W} , since not all the minimization strategies are applicable or adequate to find $\hat{\omega}$. The size of the labelling space $|\mathcal{W}| = |\mathcal{L}|^{|S|}$, discontinuities in $U(\omega)$ along \mathcal{W} or the problem of local minima, additionally all the particular of all the different minimization. Need to be taken into account while choosing the most desirable minimization strategy.

3 BUS images segmentation using optimization

In this section, the problem of delineating structures in BUS images is defined as an optimization problem that can be solved by applying the framework presented in Sect. 2. The segmentation here proposed aims at tying a label $l \in \mathcal{L}$ (*i.e.* {lesion, $\bar{\text{lesion}}$ } or {chest wall, lungs, \dots , lesion}) to each element of \mathcal{S} by simultaneously optimizing the data and pairwise terms as illustrated in Fig. 2. Choices made regarding different elements: the representation \mathcal{S} , the data term $D(\cdot)$, the pairwise term $V(\cdot)$, and the optimizer choice are summarized in Table 1 and justified thereafter (see Fig. 1 for reference).

\mathcal{S} is considered the result from an over-segmentation of the image using Quick-shift super-pixel [1]. The structures of the breast and their rendering when using a hand-held 2D US probe are sketched in Fig. 3a. Figure 3b illustrates the lexicon proposed by the ACR [7] and used by clinicians to perform their diagnosis. Thus, our aim is to generate a set of computer vision features which is able to encode the characteristic described in the lexicon. The selected features are as follows:

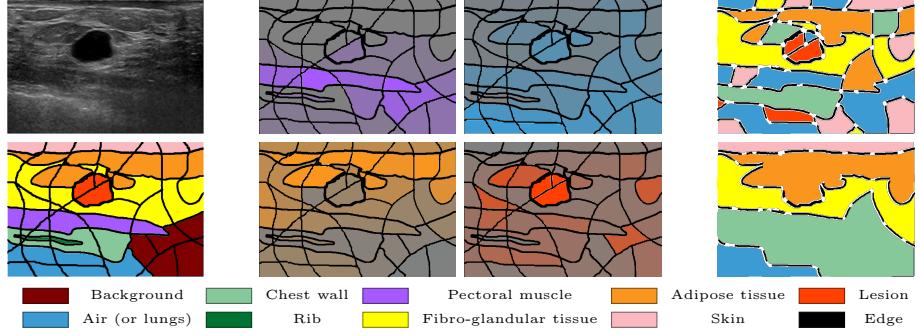


Fig. 2: Methodology Highlights. (a) BUS image example. (b) Superpixels' representation coloured using dataset's accompanying multi-label GT. (c) GT color code. (d) Data term: cost of labelling all sites as pectoral, lungs, adipose tissue or lesion. For illustration purposes, highly saturated colour indicates a low data cost - i.e., high confidence to assign the label associated with the color. (e) Pairwise term: labelling configurations with more boundaries produce higher pairwise term cost.

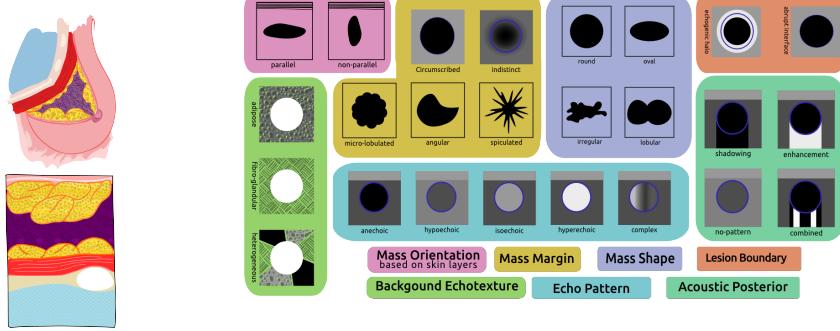


Fig. 3: Visual reference: (a) breast structures, (b) US BI-RADS lexicon

Appearance Based on the multi-labelled GT, a Median Absolute Deviation (MAD) histogram model for every tissue label is built. The Appearance feature is computed as the χ^2 distance between a histogram of s and the models generated.

Atlas Based on the multi-labelled GT, an atlas is built to encode the labels likelihood based on the location of s .

Brightness Intensity descriptors are computed based on statistics of s (*i.e.* mean, median, mode) and are compared with some intensity markers of the set S such as the minimum intensity value, the maximum, its mean, etc.

Self-Invariant Feature Transform (SIFT)-Back-of-Features (BoF) s is described as a histogram of visual words based on SIFT [6]. The dictionary is built with 36 words.

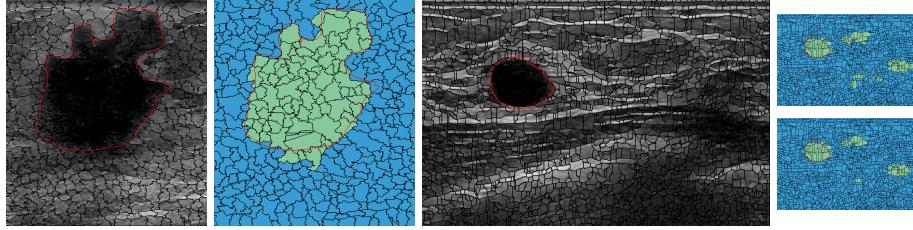


Fig. 4: Qualitative results. (a) Example 1: original image, super-pixels' delineations and GT. (b) Differences between GT and the delination resulting from super-pixels' boundary. (c) Ex. 2. (d) weak $V(\cdot, \cdot)$ (e) strong $V(\cdot, \cdot)$

The relationships between the lexicon and the descriptors described previously are depicted in Table 1. More precisely, we highlight the corresponding elements of the lexicon which is encoded by each feature. A choice regarding the encoding of the data term $D(\cdot)$ has to be made by using a ML classifier. Support Vector Machines (SVM) classifier with Radial Basis Function (RBF) kernel is selected to determine the data model during the training stage. The pairwise term in our framework was defined as in Eq. (3). The optimization method used as a solver to minimize our cost function $U(\cdot)$ is Graph-Cuts (GC). GC, where appropriately applied, allows to rapidly find a strong local minima guaranteeing that no other minimum with lower energies can be found [3]. GC is applicable if, and only if, the pairwise term favours coherent labelling configurations and penalizes labelling configurations where neighbours' labels differ such as in our case, given by Eq. (3).

4 Method evaluation and comparison

The proposed methodology is evaluated using a dataset of 16 BUS images presenting a single lesion of variable extension. The size of the lesions ranges from under 1/100 to over 1/5 of the image size. The dataset is composed of cysts, Fibro-Adenomas (FAs), Ductal Inflating Carcinomas (DICs) and Inflating Lobular Carcinomas (ILCs). Every image has accompanying multi-label GT delineating all the depicted structures. This dataset is now publicly available at <http://visor.udg.edu/dataset/#breast>

Table 1: Design choices summary

| | |
|----------------------|--|
| \mathcal{S} | Quick-Shift super-pixels |
| $D(\cdot)$ | Background Echotexture: encoded in Appearance and SIFT-BoW Echo Pattern: encoded in Appearance, Atlas and Brightness Acoustic Posterior: encoded in Atlas and Brightness |
| $V(\cdot, \cdot)$ | homogeneity as Eq. (3) |
| $\arg \min U(\cdot)$ | Graph-Cuts |

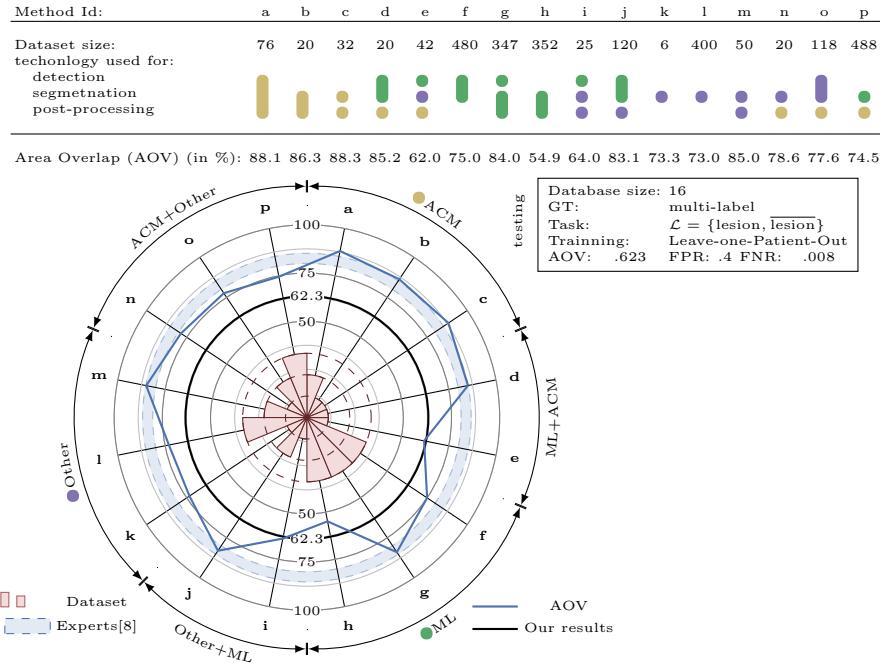


Fig. 5: Quantitative results compilation and comparison

The lack of publicly available data and source code, limits the comparison between the different methods. For this study, the results published by the other authors have been collected and expressed in terms of AOV in order to share a common metric. Further details can be found in [5] and summarized in Fig. 5.

Figure 5 is divided into three main parts: (i) a table on the top summarizes the core stages of each study framework, (ii) a legend box on the right side informs our testing setup and, (iii) a comparison of the different metrics in a radial manner. An extra element is also represented in this radial representation: a blue swatch delimited by two blue dashed lines. The boundaries of this swatch correspond to the performance of some expert radiologists based on an inter- and intra-observer experiments carried out by Pons et al. [8]. It is interesting to note that some methodologies outperform this swatch. A publicly available dataset should allow a better comparison in that regard.

The results point out the inherent capabilities of ML to cope with data scalability and variability, induce its usage in conjunction with larger datasets. Whereas, Active Contour Model (ACM) methodologies show its effectiveness to model the boundary in a natural manner.

For our proposed framework, the performance in terms of AOV lies within the state-of-the-art despite its final delineation limited by the capacity of the superpixels to snap the desired boundary. Figure 4 shows some qualitative results

where there are limitations of labeling super-pixels when compared with hand-drawn GT. Figure 4 also illustrates the influence of the pair-wise term.

5 Conclusions

This work presents a segmentation strategy to delineate lesions in BUS images using an optimization framework that takes advantage of all the facilities available when using ML techniques. Despite the limitation that the final segmentation is subject to the super-pixels' boundaries, the AOV results reported here are similar to those reported by other methodologies in the literature. A higher AOV result can be achieved by refining the delineation resulting from our proposed framework by post-processing it with an ACM. In this manner, the contour constraints could be applied to achieve a more natural delineation.

References

1. Achanta, R., et al.: SLIC superpixels compared to state-of-the-art superpixel methods (2012)
2. Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International journal of computer vision* 72(2) (2007)
3. Delong, A., Osokin, A., Isack, H.N., Boykov, Y.: Fast approximate energy minimization with label costs. *International Journal of Computer Vision* 96(1), 1–27 (2012)
4. Jemal, A., et al.: Global cancer statistics. *CA: A Cancer Journal for Clinicians* 61 (2011)
5. Massich, J.: Deformable object segmentation in ultra-sound images. Ph.D. thesis
6. Massich, J., et al.: Sift texture description for understanding breast ultrasound images. In: Breast Imaging, Lecture Notes in Computer Science, vol. 8539, pp. 681–688. Springer International Publishing (2014)
7. Mendelson, E., Baum, J., WA, B., et al.: BI-RADS: Ultrasound, 1st edition in: D'Orsi CJ, Mendelson EB, Ikeda DM, et al: Breast Imaging Reporting and Data System: ACR BIRADS – Breast Imaging Atlas. American College of Radiology (2003)
8. Pons, G., Martí, J., Martí, R., Ganau, S., Vilanova, J., Noble, J.: Evaluating lesion segmentation in breast ultrasound images related to lesion typology. *Journal of Ultrasound in Medicine* (2013)
9. Smith, R.A., et al.: American cancer society guidelines for breast cancer screening: update 2003. *CA: a cancer journal for clinicians* 53(3), 141–169 (2003)
10. Stavros, A.T., Thickman, D., Rapp, C.L., Dennis, M.A., Parker, S.H., Sisney, G.A.: Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions. *Radiology* 196(1), 123–34 (1995)



Classifying DME vs Normal SD-OCT volumes: A review

Joan Massich, Mojdeh Rastgoo, Guillaume Lemaître, Carol Cheung, Tien Wong, Désiré Sidibé, Fabrice Mériadeau

► To cite this version:

Joan Massich, Mojdeh Rastgoo, Guillaume Lemaître, Carol Cheung, Tien Wong, et al.. Classifying DME vs Normal SD-OCT volumes: A review. 23rd International Conference on Pattern Recognition, Dec 2016, Cancun, Mexico. <hal-01376469>

HAL Id: hal-01376469

<https://hal-univ-bourgogne.archives-ouvertes.fr/hal-01376469>

Submitted on 4 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Classifying DME vs Normal SD-OCT volumes: A review

Joan Massich^{*¶||}, Mojdeh Rastgoo^{*¶}, Guillaume Lemaître^{*¶}, Carol Y. Cheung[†],
Tien Y. Wong[†], Désiré Sidibé^{*}, Fabrice Mériadeau^{*‡}

^{*}LE2I UMR6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté,
12 rue de la Fonderie, 71200 Le Creusot, France

[†]Singapore Eye Research Institute, Singapore National Eye Center, Singapore

[‡]Centre for Intelligent Signal and Imaging Research (CISIR), Electrical & Electronic Engineering Department,
Universiti Teknologi Petronas, 32610 Seri Iskandar, Perak, Malaysia

[¶]Indicates equal contribution. Ordering determined by Rock-Paper-Scissors-Lizard-Spock.

^{||}Corresponding author: joan.massich@u-bourgogne.fr

Abstract—This article reviews the current state of automatic classification methodologies to identify Diabetic Macular Edema (DME) versus normal subjects based on Spectral Domain OCT (SD-OCT) data. Addressing this classification problem has valuable interest since early detection and treatment of DME play a major role to prevent eye adverse effects such as blindness.

The main contribution of this article is to cover the lack of a public dataset and benchmark suited for classifying DME and normal SD-OCT volumes, providing our own implementation of the most relevant methodologies in the literature. Subsequently, 6 different methods were implemented and evaluated using this common benchmark and dataset to produce reliable comparison.

Index Terms—Diabetic Macular Edema (DME), Spectral Domain OCT (SD-OCT), Machine Learning (ML), benchmark,

I. INTRODUCTION

Diabetic Retinopathy (DR), and more particularly Diabetic Macular Edema (DME), are leading causes of irreversible vision loss and the most common eye diseases in individuals with diabetes. Taking into account that the number of individuals affected by diabetes diseases are expected to grow exponentially in the next decade [1], developing methodologies for early detection and treatment of DR and DME has become a priority to prevent adverse effects.

The main focus of this work is to describe the actual state of DME detection in Optical Coherence Tomography (OCT) images. DME presents an increase in retinal thickness within 1 disk diameter of the fovea center with or without hard exudates and sometimes associated with cysts [2]. Spectral Domain OCT (SD-OCT) is an emerging eye imaging modality providing cross-sectional retinal morphology information [3], which cannot be estimated from more established eye imaging modalities such as fundus imaging.

The initial efforts of the ophthalmic community in developing technologies for SD-OCT have been placed in segmenting the retinal layers, which is a necessary step for retinal thickness measurements [4, 5]. However, latter efforts address the specific problem of DME automatic detection in OCT volumes. These efforts reveal the needs to address: (i) enhancing the quality of OCT volumes, (ii) finding pathology signs, and (iii) appropriate classification strategies.

Advances in any of those regards is of great interest since (i) manual evaluation of SD-OCT volumetric scans is expensive and time consuming [6]; (ii) SD-OCT acquisition has some shortcomings due to eye movements during the scanning [7], reflectivity nature of the retina [8], high level of noise and inconsistent quality of the images; (iii) due to the coexistence of multiple pathologies [7] as well as large intra-pathology variability, consistently identifying pathology-specific biomarkers remains challenging [6].

The rest of this article is structured as follows: Section II offers a general idea of the literature state-of-the-art in SD-OCT volume classification. Section III reviews some publicly available datasets and states the need for another one that suits the classification task here described. Section IV proposes an experimental benchmark to compare different methodologies presented in Sect. II. Section V reports and discusses the obtained results, while Sect. VI wraps up our thoughts regarding this work and its possible direction.

II. BACKGROUND

This section reviews works straightly addressing the problem of classifying OCT volumes as normal or abnormal, regardless of the targeted pathology. The methods are categorized in terms of their learning strategy, namely supervised or semi-supervised learning.

A. Supervised methods

Supervised learning is based on a fully annotated and labeled training set. In this approach, the labeled training data are used to train the classifier function later used for prediction. Figure 1 illustrates a prevalent framework for supervised learning. Each SD-OCT volume undergoes: (i) *pre-processing* to reduce noise and other acquisition deficiencies which alter the images; (ii) *feature detection* to quantify visual cues like appearance, texture, shape, etc.; (iii) *mapping* in which a sample is either considered as whole (i.e., global) or partitioned into a set of sub-elements (i.e., local dense/sparse patches, pyramid, etc.); (iv) *feature representation* to associate a descriptor (e.g., concatenation, statistics, histogram,

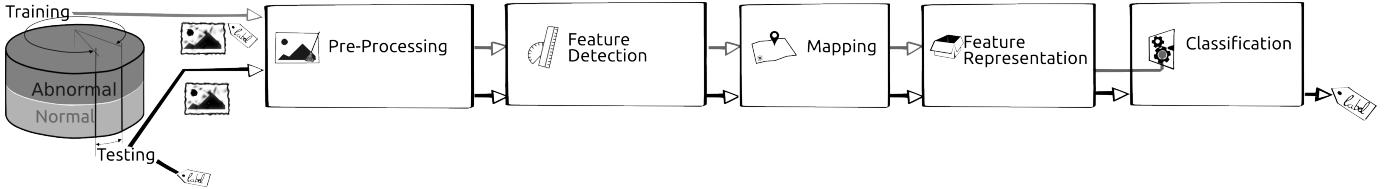


Fig. 1. Common framework

Principal Component Analysis (PCA), Bag-of-Words (BoW), etc.) for each element from the *mapping-stage*. This descriptor packages the visual cues related to the sample; (v) *classification* to determine the associated class of each sample.

Venhuizen *et al.* propose a classification method to distinguish between Age-related Macular Degeneration (AMD) and normal SD-OCT volumes using BoW models [6]. A set of keypoints are detected and selected at each individual B-scan, by keeping the salient points included in the top 3% of the vertical gradient values. Around each of these keypoints, a $9 \text{ px} \times 9 \text{ px}$ texton is extracted, generating a feature vector of 81 dimensions, later reduce to 9 using PCA. All extracted feature vectors are used to create a codebook using k -means clustering. Then, each OCT volume descriptor is represented as a histogram that captures the codebook occurrences and are classified by a Random Forest (RF) composed of 100 trees. The method is tested using a publicly available dataset of 384 OCT volumes [9], achieving an Area Under the Curve (AUC) of 0.984.

Srinivasan *et al.* propose a classification method to distinguish DME, AMD, and normal SD-OCT volumes [10]. Each OCT slice is pre-processed using Block Matching 3D filtering (BM3D) to reduce the speckle noise and is flattened to reduce the inter-patient retinal curvature variations. A multi-resolution pyramid is generated for each pre-processed slice and a Histogram of Oriented Gradients (HOG) feature is computed for each layer. These features are classified using a linear Support Vector Machines (SVM). Note that each individual B-scan is classified into one of the three categories, namely DME, AMD, and normal, and a volume is label to a given class by taking the majority vote of all B-scans. This method is also tested using a publicly available dataset, composed of 45 patients equally subdivided into the three targeted classes. Correct classification rates of 100%, 100% and 86.67% are obtained for normal, DME, and AMD patients, respectively.

Extending the previous work, Alsaih *et al.* aggregate Local Binary Patterns (LBP) to HOG in order to add texture information and reduce the number of dimension using PCA [11].

Lemaître *et al.* propose a method based on LBP features to describe the texture of OCT images and dictionary learning using the BoW models [12]. In this method, the OCT images are first pre-processed using Non-Local Means (NLM) filtering, to reduce the speckle noise. Then, the volumes are mapped into a discrete set of structures: (i) local corresponding to patches, or (ii) global corresponding to volume slices or

the whole volume. According to the chosen mapping, LBP or LBP from Three Orthogonal Planes (LBP-TOP) texture features are extracted and represent each volume through histogram, PCA, or BoW representation. The final feature descriptors are classified using RF classifier. This methodology is tested against Venhuizen *et al.* [6] using public and non-public datasets showing an improvement within the results by achieving a Sensitivity (SE) of 87.5% and a Specificity (SP) of 75%.

Liu *et al.* propose a methodology aiming at classifying B-scan rather than volume. The classification goal is to distinguish between macular pathology and normal OCT B-scan images using LBP and gradient information as attributes [7]. Each OCT slice is flattened before to create a 3-level multiti-scale spatial pyramid. From each layer of this pyramid, edges are extracted and LBP descriptors are computed for the flattened slice and the edge map. All the obtained histograms are concatenated into a global descriptor whose dimensions are reduced using PCA. Finally, a SVM with a Radial Basis Function (RBF) kernel is used as classifier. A detection rate with an AUC of 0.93 is achieved, using a dataset of 326 OCT scans with various pathologies.

Albarraak *et al.* propose another classification framework to differentiate AMD and normal volumes [13]. Each OCT slice undergoes two pre-processing routines: (i) a joint denoising and cropping step using the split Bregman isotropic total variation algorithm and (ii) a flattening step by fitting a second-order polynomial using a least-square approach. Then, LBP-TOP and HOG combined with LBP-TOP features are extracted from individual sub-volumes from each original cropped volume. These features are concatenated into a single feature vector per OCT volume and its dimension is reduced using PCA. Finally, a Bayesian network classifier is used to classify the volumes. The classification performance of the framework in terms of SE and SP achieves 92.4% and 90.5%, respectively, outperforming the method of Liu *et al.* [7], using a dataset composed of 140 OCT volumes.

Anantrasirichai *et al.* propose to detect glaucoma in OCT images based on a variety of texture descriptor [14]. The texture information is described through LBP, Gray-level co-occurrence matrix (GLCM), wavelet, granulometry, run length measures, and intensity level distributions in combination with retinal layer thickness estimation, without any pre-processing. Each feature vector is projected using PCA before to be classified using an SVM with both linear and RBF kernel. Testing with rather a small dataset of 24 OCT volumes, their

proposed method achieves an Accuracy (ACC) of 85 % while using layer thickness and textural informations.

B. Semi-supervised methods

Sankar *et al.* propose to use a semi-supervised strategy to classify DME vs. normal OCT volumes based on appearance modeling of normal OCT images using Gaussian Mixture Model (GMM) [15]. The main difference between this method and the supervised methodologies lies in the fact that only normal volumes are used to train the system.

For each OCT volume, the B-scans are denoised using NLM filtering, flattened, and resized to ensure homogeneous dimension across all volumes. Each B-scan is vectorized and projected into a lower-dimensional space with p dimensions using PCA. Subsequently, normal B-scans are modelled using a GMM in which the number of mixture components K is determined on a validation set. At the testing stage, a scan is classified as normal or DME depending of its Mahalanobis distance to the learnt model; if the distance is greater than the 97.5% quantile of the Chi-squared distribution with p degree of freedom. Therefore, a volume is classified as abnormal if the number of abnormal slice is greater than a given threshold, previously determined during the validation procedure. A SE and SP of 93.8% and 80.0% are respectively achieved on a cohort of 32 patients.

III. DATA

Proper comparison of different methodologies require a common dataset to test these methodologies. The lack of public data for comparing methodologies is recurrent claim in the medical image community [17]. To amend this limitation, Duke University made two SD-OCT datasets available to the ophthalmic community [9, 10].

The rest of this section highlights the advantages and disadvantages of the datasets provided by Duke University, points out why this datasets cannot be used for our purposes and finally describes our data: the *SERI dataset* [18].

The former dataset from Duke University consists of 384 OCT annotated volumes classified either as AMD or normal cases. Despite the advantage of testing in large datasets, this dataset cannot be used to conduct our study since we are interested in DME and not AMD. This dataset has been used by Venhuizen *et al.* [6] to test their method since their main interest is AMD detection. The later dataset from Duke University consists of 45 pre-processed OCT volumes and labeled as AMD, DME, and normal. Despite this dataset is suitable to our goal of classifying DME vs normal volumes, the dataset has been dropped since there is no access to the original data. All volumes have been denoised, aligned and cropped. This dataset has been used to conduct the experimentation reported by Srinivasan *et al.* [10].

The dataset to conduct our study has been acquired by the SERI, using CIRRUS TM (Carl Zeiss Meditec, Inc., Dublin, CA) SD-OCT device. This dataset consists of 32 OCT volumes, subdivided into 16 DME and 16 normal cases. Each volume contains 128 B-scans with a resolution of

512 px \times 1,024 px. All SD-OCT images have been read and assessed by trained graders and identified as normal or DME cases, based on evaluation of retinal thickening, hard exudates, intraretinal cystoid space formation and subretinal fluid (see Fig. 2).

IV. EXPERIMENTAL SETUP

The experimental set-up is summarized in Table I, where the most relevant works in Sect. II are formulated as the 5-steps standard classification procedure described in Fig. 1.

A. Implementation details

The experiments, described in this work, are publicly available at [19] allowing for further comparisons and improvements. All the methods in Table I have been developed using *protoclass* [20], a rapid prototyping toolkit to perform image processing and Machine Learning (ML) tasks. Furthermore, each method has been implemented as a plug-in to [19], so that all methods can be evaluated in a common framework¹.

Note that Liu *et al.* train the algorithm at the B-scan level, and SERI dataset provides Ground Truth (GT) at volume level only. Thus, two strategies have been explored to solve this issue: (i) similarly to Srinivasan *et al.*, at training stage, all B-scans are considered as abnormal for a DME volume and at testing stage, a majority vote rule is applied to whether label a volume as abnormal or not; (ii) similarly to Venhuizen *et al.*, an approach using BoW is used. From the methods reviewed in Sect. II, we decline to implement Albarak *et al.* and Anantrasirichai *et al.*. The former do not provide sufficient implementation details to replicate their results [13]; while, the latter use a descriptor based on the layer thickness which require a layer segmentation stage using a generic segmentation algorithm and further user validation [14].

B. Evaluation

All the experiments are evaluated in terms of SE and SP (see Fig. 3) using the Leave-Two-Patient Out Cross-Validation (LTPO-CV) strategy.² SE evaluates the performance of the classifier with respect to the positive class, while the SP does the same with respect to negative class. LTPO-CV keeps two volumes (one normal and one DME) for testing while the remaining volumes are used as training. The advantage of using LTPO-CV over Leave-One-Patient Out Cross-Validation (LOPO-CV) is that LTPO-CV keeps the training data balanced. The main drawback of using LTPO-CV (or LOPO-CV) is that despite reporting robust performance estimators, variance of these descriptors cannot be computed. Despite this limitation, LTPO-CV strategy has been adopted here, since regular Computer Vision (CV) cannot be applied due to the reduced size of the dataset.

¹See table I for standalone repositories of each method. All repositories provide tests to ensure that our implementation comply with the original work.

²the same evaluation strategy was applied [12]

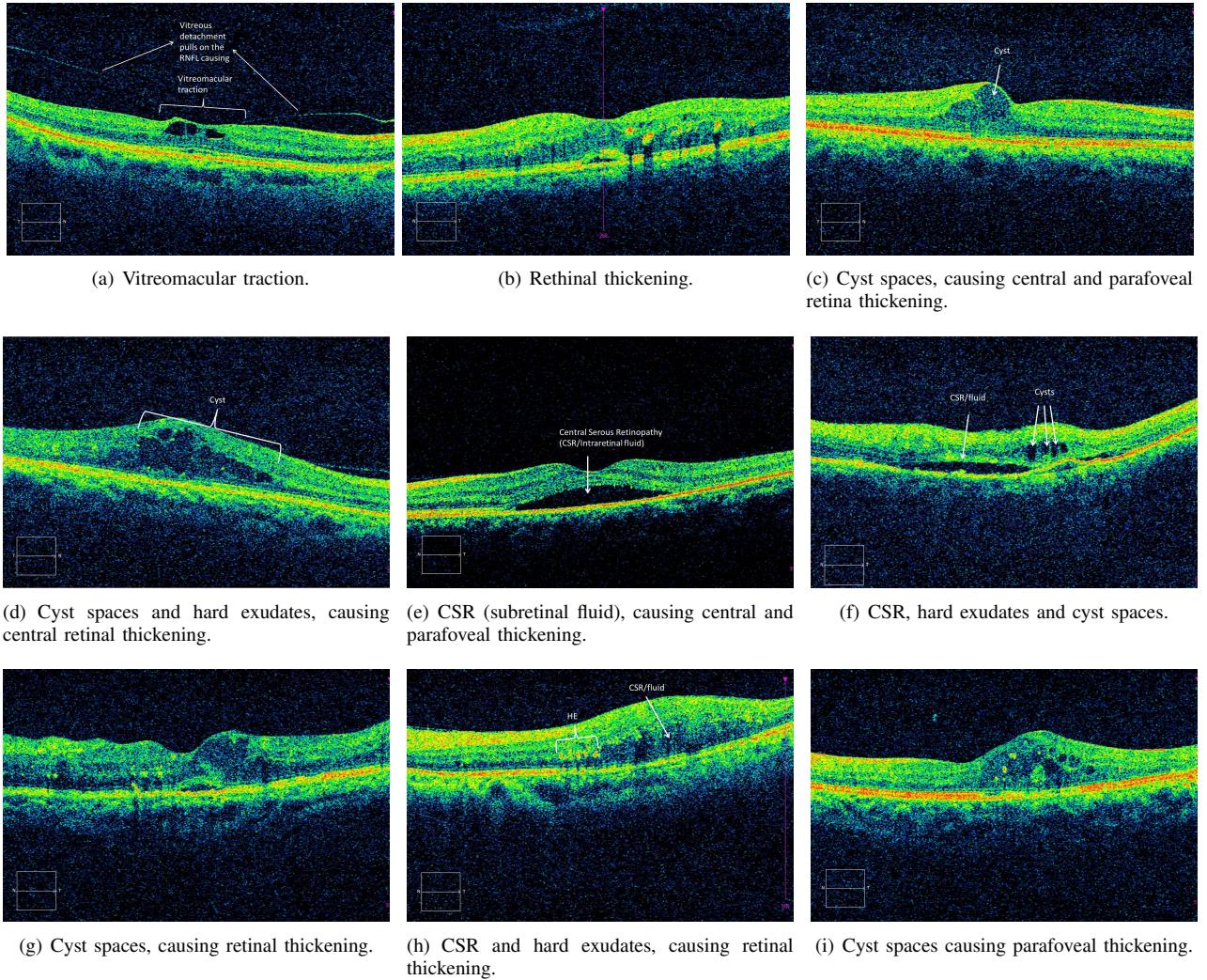


Fig. 2. Examples of DME cases in Singapore Eye Research Institute (SERI) dataset.

V. RESULTS AND DISCUSSION

The entire set of experiments with their associated results can be found in [19], while Table I shows the configuration leading to the best results of each method. The results are reported in terms of SE and SP (see Sect. IV-B).

Lemaitre *et al.* achieve the best results when using LBP-TOP features, a global mapping, and histogram representation [12]. Alsahai *et al.* perform better when using HOG features with PCA representation [16]. Our interpretation of Liu *et al.*, as proposed in Sect. IV, achieves the best results when using majority voting instead of BoW models. Refer to Table I for configuration details of the remaining methods.

Results in [19] indicate two main findings with major impact: (i) features describing the entire volume rather than each B-scan are more discriminative; and (ii) a pre-processing stage with denoising is fundamental.

Other observations include the facts that (i) to represent B-scans, local mapping in conjunction with dimension reduction, either using PCA or BoW, improve the results. However, the

combination of both decreases the performance in comparison to non reduced histogram representation; (ii) building BoW models from the concatenation of all features for each B-scan, might lead to the curse of dimensionality since 128 samples per volume is not enough to describe a space with a number of dimensions of the order of thousands; which could explain the over-fitting using RBF-SVM as in [21].

VI. CONCLUSION AND FURTHER WORK

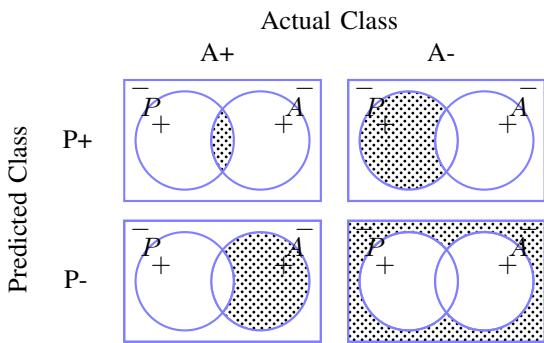
The work here presented states the relevance of developing methodologies to automatically differentiate DME *vs.* normal SD-OCT scans. This article offers an overview of the state-of-the-art of DME detection and provides a public benchmarking to facilitate further studies. In this regard, there are two crucial aspects to improve the work here presented: (i) enlarge the dataset. (ii) reach out to other authors in order to enlarge this benchmark with additional methods and improve the existing approaches.

TABLE I
CORRESPONDENCE BETWEEN THE MOST RELEVANT METHODOLOGIES REVIEWED IN SECT. II AND THE PROPOSED EXPERIMENTAL FRAMEWORK.

| Ref | Pre-processing | Features | Mapping | Representation | Classification |
|-----------------------------------|--|----------------------|-----------------|-------------------------|------------------------------------|
| Venuhuzen <i>et al.</i> [6, 22] | — | Texton | Local | PCA BoW | RF |
| Srinivasan <i>et al.</i> [10, 23] | Denoising (BM3D) Flattening Cropping | HOG | Global | — | Linear-SVM |
| Lemaître <i>et al.</i> [12, 24] | Denoising (NLM) | LBP LBP-TOP | Local Global | PCA BoW Histogram | RF |
| Alsaih <i>et al.</i> [11, 16] | Denoising (BM3D) Flattening Cropping | LBP HOG | Local | PCA Histogram | Linear-SVM |
| Liu <i>et al.</i> [7, 21] | Flatten Aligned | Edge LBP | Local | PCA BoW | RBF-SVM |
| Sankar <i>et al.</i> [15, 25] | Denoising (NLM) Flattening Cropping | Pixel intensities | Global | PCA | Mahalanobis -distance to GMM |

TABLE II
SUMMARY OF THE CLASSIFICATION PERFORMANCE IN TERMS OF SE AND SP IN (%).

| Lemaitre <i>et al.</i> [12] | Sankar <i>et al.</i> [15] | Alsaih <i>et al.</i> [16] | Srinivasan <i>et al.</i> [10] | Liu <i>et al.</i> [7] | Venuhuzen <i>et al.</i> [6] |
|-----------------------------|---------------------------|---------------------------|-------------------------------|-----------------------|-----------------------------|
| SE | 87.5 | 81.3 | 75.0 | 68.8 | 68.8 |
| SP | 75.0 | 62.5 | 87.5 | 93.8 | 58.8 |



(a) Confusion matrix with truly and falsely positive samples detected (TP, FP) in the first row, from left to right and the falsely and truly negative samples detected (FN, TN) in the second row, from left to right.

$$SE = \frac{TP}{TP+FN} \quad SP = \frac{TN}{TN+FP}$$

(b) SE and SP evaluation, corresponding to the ratio of the doted area over the blue area.

Fig. 3. Evaluation metrics: (a) confusion matrix, (b) SE - SP

REFERENCES

- [1] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, “Global prevalence of diabetes estimates for the year 2000 and projections for 2030,” *Diabetes care*, vol. 27, no. 5, pp. 1047–1053, 2004.
- [2] Early Treatment Diabetic Retinopathy Study Group, “Photocoagulation for diabetic macular edema: early treatment diabetic retinopathy study report no 1,” *JAMA Ophthalmology*, vol. 103, no. 12, pp. 1796–1806, 1985.
- [3] Y. T. Wang, M. Tadarati, Y. Wolfson, S. B. Bressler, and N. M. Bressler, “Comparison of Prevalence of Diabetic Macular Edema Based on Monocular Fundus Photography vs Optical Coherence Tomography,” *JAMA Ophthalmology*, pp. 1–7, Dec 2015.
- [4] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, “Automatic segmentation of seven retinal layers in sd-oct images congruent with expert manual segmentation,” *Optic Express*, vol. 18, no. 18, pp. 19413–19428, 2010.
- [5] R. Kafieh, H. Rabbani, M. D. Abramoff, and M. Sonka, “Intra-retinal layer segmentation of 3d optical coherence tomography using coarse grained diffusion map,” *Medical Image Analysis*, vol. 17, pp. 907–928, 2013.
- [6] F. G. Venuhuzen, B. van Ginneken, B. Bloemen, M. J. P. P. van Grisven, R. Philipsen, C. Hoyng, T. Theelen, and C. I. Sanchez, “Automated age-related macular degeneration classification in OCT using unsupervised feature learning,” in *SPIE Medical Imaging*, vol. 9414, 2015, p. 941411.
- [7] Y.-Y. Liu, M. Chen, H. Ishikawa, G. Wollstein, J. S. Schuman, and R. J. M., “Automated macular pathology diagnosis in retinal oct images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding,” *Medical Image Analysis*, vol. 15, pp. 748–759, 2011.
- [8] J. S. Schuman, C. A. Puliafito, J. G. Fujimoto, and J. S. Duker, *Optical coherence tomography of ocular diseases*. SLACK incorporated, 2004.
- [9] S. Farsiu, S. J. Chiu, R. V. O’Connell, F. A. Folgar, E. Yuan, J. A. Izatt, C. A. Toth, A.-R. E. D. S. . A. S. D. O. C. T. S. Group *et al.*, “Quantitative classification of eyes with and without intermediate

- age-related macular degeneration using optical coherence tomography," *Ophthalmology*, vol. 121, no. 1, pp. 162–172, 2014.
- [10] P. P. Srinivasan, L. A. Kim, P. S. Mettu, S. W. Cousins, G. M. Comer, J. A. Izatt, and S. Farsiu, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomedical Optical Express*, vol. 5, no. 10, pp. 3568–3577, 2014.
- [11] K. Alsaih, G. Lemaitre, J. Massich, M. Rastgoo, D. Sidibé, T. Y. Wong, E. Lamoureux, D. Milea, C. Y. Cheung, and F. Mériauadeau, "Classification of sd-oct volumes with multi pyramids, lbp and hog descriptors: application to dme detections," in *38th IEEE Engineering in Medicine and Biology Society (EMBC)*. Orlando, USA: IEEE, 2016.
- [12] G. Lemaitre, M. Rastgoo, J. Massich, S. Sankar, F. Mériauadeau, and D. Sidibe, "Classification of SD-OCT volumes with LBP: Application to dme detection," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI), Ophthalmic Medical Image Analysis Workshop (OMIA)*, 2015.
- [13] A. Albarak, F. Coenen, and Y. Zheng, "Age-related macular degeneration identification in volumetric optical coherence tomography using decomposition and local feature extraction," in *Proceedings of 2013 International Conference on Medical Image, Understanding and Analysis*, 2013, pp. 59–64.
- [14] N. Anantrasirichai, A. Achim, J. E. Morgan, I. Erchova, and L. Nicholson, "Svm-based texture classification in optical coherence tomography," in *IEEE 10th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2013, pp. 1332–1335.
- [15] S. Sankar, D. Sidibé, Y. Cheung, T. Wong, E. Lamoureux, D. Milea, and F. Mériauadeau, "Classification of sd-oct volumes for dme detection: an anomaly detection approach," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2016, pp. 97852O–97852O.
- [16] K. Alsaih, G. Lemaitre, J. Massich, M. Rastgoo, D. Sidibe, and F. Mériauadeau, "Classification of SD-OCT volumes with multi pyramids, LBP and HOG descriptors: application to DME detections," Aug. 2016. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.50145>
- [17] M. L. Giger, H.-P. Chan, and J. Boone, "Anniversary paper: History and status of CAD and quantitative image analysis: the role of medical physics and AAPM," *Medical physics*, vol. 35, no. 12, p. 5799, 2008.
- [18] G. Lemaitre, J. Massich, M. Rastgoo, and F. Mériauadeau, "Seri dataset," Sep. 2016. [Online]. Available: <http://i2cvb.github.io/#retino>
- [19] G. Lemaitre, J. Massich, and M. Rastgoo, "I2Cvb Benchmark: automatic classification of DME vs normal SD-OCT volumes," Apr. 2016. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.61976>
- [20] G. Lemaitre, "Protoclass, a rapid prototyping tool for image processing and machine learning," Apr. 2016. [Online]. Available: <http://glemaître.github.io/protoclass/>
- [21] G. Lemaitre, J. Massich, and M. Rastgoo, "liu-2011-oct: ICPR2016," Apr. 2016. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.50140>
- [22] ———, "venhuizen-2015-feb: Icpr2016," Feb. 2016. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.61981>
- [23] G. Lemaitre, "srinivasan-2014-oct: Icpr 2016," Apr. 2016. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.50136>
- [24] G. Lemaitre, J. Massich, and M. Rastgoo, "retinopathy: Jo-omia-2015," Nov. 2015. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.34277>
- [25] G. Lemaitre, J. Massich, M. Rastgoo, S. Sankar, D. Sidibe, and F. Mériauadeau, "sankar-2016-feb: Icpr2016," Feb. 2016. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.50137>