UNIVERSIDAD POLITÉCNICA DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIEROS
DE TELECOMUNICACIÓN

Image processing methods for computer-aided screening for disease

PhD. Thesis

Author

Juan José Gómez Valverde

Department of Electronic Engineering

2018

Biomedical Image Technologies Laboratory
Department of Electronic Engineering
Escuela Técnica Superior de Ingenieros de Telecomunicación
Universidad Politécnica de Madrid

Ph.D. Thesis

Image processing methods for computer-aided screening for disease

Author

Juan José Gómez Valverde

Advisors

María Jesús Ledesma Carbayo and Andrés Santos Lleó

Copyright © 2018 Juan José Gómez Valverde

Abstract

Screening for disease has been a widely accepted practice in healthcare since the twentieth century. Nowadays, in most developed countries we can find organized screening programs in different types of cancer (breast, cervical and colorectal are the most common) and for antenatal and newborns. Two clinical fields where the potential benefits of screening are very promising and have received worldwide attention are eye diseases, like diabetic retinopathy or glaucoma, and skin diseases, like melanoma and non-melanoma skin cancer.

Screening programs have adopted a broad range of technological facilities. In this work we get an extensive insight in the role of image processing, telemedicine and Computer-Aided-Detection (CAD) systems and the benefits that can provide to aid in the fulfillment of the good practices and protocols established for a successful service. The use of medical images has become an indispensable diagnostic tool in many current screening programs. The great advance of medical imaging and image processing, their intensive use in the majority of the specialties to aid in the detection and follow-up of most conditions and their availability in health provider centers, have driven to be one of the tests typically performed in a screening campaign. Among all imaging modalities present in current screenings, we will focus on image processing methods for Color Fundus (CF) and Optical Coherence Tomography (OCT) images. While CF is basically used in retinal imaging, OCT can be used in retinal and skin imaging among many other fields.

Telemedicine and e-health technology have played a significant role assisting in the provision of screening services. Associated to the most established screening implementations there have been remarkable telemedicine initiatives. The intensive use of imaging for the diagnosis and monitoring of eye diseases make telemedicine ideal for tele-screening, but it is still an open research field in pathologies like glaucoma that demands further developments to demonstrate its potential benefits. Glaucoma is one of the leading causes of global irreversible vision loss and it is estimated that affects 70 million people worldwide. This pathology represents a promising opportunity for screening due to the relevant percentage, up to 50% of population suffering the disease that is not aware of it and the benefits that an early diagnosis and treatment can represent to the patient. In this context, we have designed and implemented a new tele-screening tool with a multi-specialty and multi-disease approach to support some of the challenges identified in current screening services. Besides, the tool was tested with good results, on more than 1000 patients in an ophthalmology screening campaign specially focused on glaucoma.

This Thesis has also studied CAD systems which have demonstrated in many medical applications that can assist physicians in the interpretation of images for the detection and diagnosis of multiple pathologies. From all the steps involved in a typical CAD system we paid special attention to two of them: the pre-processing for image enhancement and the automatic classification of a specific disease. In the pre-processing we present

a new imaging enhancement method for OCT images that can serve as the first step in a CAD system for skin imaging as well as to improve emerging portable prototypes to be used in screening programs. We have tested the method successfully in several skin datasets, including an innovative qualitative assessment methodology to evaluate the clinical value of this type of image processing. Finally, for the automatic classification part, we present an extensive study of the application of deep convolutional neural networks to the challenging task of detecting glaucoma in color fundus images, which could serve as a valuable tool to be incorporated in CAD systems for large glaucoma screening campaigns.

In this thesis we have examined, implemented and tested new advanced methods in several relevant technological components present in many screening services. The results reveal new promising lines of work that we hope to continue exploring in the future.

El cribado de enfermedades es una práctica muy extendida en los servicios de salud desde el siglo XX. En nuestros días, en la mayoría de los países desarrollados podemos encontrar programas de cribado en fetos y recién nacidos y diferentes tipos de cáncer, siendo los más extendidos el de mama, de cérvix y colorrectal. Dos especialidades médicas donde los potenciales beneficios del cribado son muy prometedores y han recibido gran repercusión en los últimos tiempos, son la oftalmología, en patologías como la retinopatía diabética o el glaucoma, y en dermatología, en enfermedades como el cáncer de piel melanoma y no melanoma.

Los programas de cribado han adoptado una gran variedad de soluciones tecnológicas para cumplir sus exigentes requerimientos. En este trabajo profundizaremos sobre todo en el papel del procesamiento de imágenes biomédicas, la telemedicina y los sistemas de ayuda al diagnóstico, y los beneficios que pueden proporcionar en el cumplimiento de los protocolos y buenas prácticas reconocidas para ofrecer un servicio de cribado de calidad. El uso de imágenes médicas se ha convertido en una herramienta indispensable en muchos programas de cribado. Los avances en nuevos dispositivos de imagen y procesamiento, su utilización intensiva en la mayoría de las especialidades para ayudar en la detección y seguimiento de múltiples patologías y su disponibilidad en los centros de salud han facilitado y potenciado su uso como una de las pruebas más habituales en una campaña de cribado. En este trabajo nos centraremos en imágenes de fondo de ojo en color y en imágenes de tomografía de coherencia óptica. Mientras que las primeras se utilizan fundamentalmente para visualizar la retina, la segunda tiene aplicaciones en innumerables especialidades, entre ellas para el estudio de enfermedades de la retina o de la piel.

La telemedicina ha jugado un rol muy relevante en la provisión y soporte de servicios de cribado. Podemos encontrar numerosas iniciativas basadas en telemedicina vinculadas a programas de cribado ampliamente implantados. El uso intensivo de imágenes en el diagnóstico y seguimiento de enfermedades oculares hace que la telemedicina sea especialmente adecuada en este campo. Pero aún hoy es un campo abierto a la investigación en patologías como el glaucoma, donde se demandan nuevos desarrollos e implementaciones que ayuden a demostrar sus potenciales beneficios. El glaucoma es una de las causas más habituales de pérdida de visión y se estima que la sufren en el mundo 70 millones de personas. Es una patología que representa una prometedora oportunidad para el cribado debido a que se estima que hasta el 50% de las personas afectadas lo desconocen y que la detección temprana y el tratamiento en las primeras etapas de la enfermedad pueden ser muy beneficiosos para el paciente. En este contexto, presentamos el diseño e implementación de una nueva herramienta de tele-cribado con un enfoque multi-especialidad y multi-enfermedad para solucionar algunos de los retos identificados actuales y futuros de los servicios de cribado. Adicionalmente, la herramienta ha sido probada

intensivamente en el contexto de un ambicioso programa de detección de enfermedades oftalmológicas con gran relevancia en el glaucoma y que ha atendido a más de 1000 pacientes.

Esta Tesis ha estudiado también los sistemas de ayuda a la detección que han demostrado en múltiples aplicaciones médicas que pueden asistir a especialistas médicos en la interpretación de imágenes para la detección y diagnóstico de innumerables enfermedades. De todos los elementos que forman parte de estos sistemas nos centramos en dos: el procesamiento para la mejora de imagen y la clasificación automática de una patología concreta, el glaucoma. Proponemos un nuevo método de mejora de imágenes de tomografía de coherencia óptica que puede servir como primer paso en sistemas de ayuda a la detección de imágenes de piel y para mejorar los nuevos dispositivos compactos y portables que se empiezan a probar en servicios de cribado. Hemos probado de forma exitosa el método propuesto en varios conjuntos de datos de imagen, incluyendo una nueva metodología de evaluación cualitativa que ayude a conocer el valor clínico de este tipo de procesamientos. Finalmente, para la parte de clasificación automática, describimos un extenso estudio de la aplicación de redes neuronales profundas para la compleja tarea de detectar glaucoma en imágenes de fondo de ojo, que puede servir de valiosa herramienta en campañas de cribado de glaucoma a gran escala.

En este trabajo hemos examinado, implementado y probado nuevos métodos y herramientas en varios componentes tecnológicos relevantes presentes en muchos servicios de cribado actuales. Los resultados muestran nuevas prometedoras líneas de trabajo que esperamos seguir explorando en el futuro.

A mi padre, siempre en mi recuerdo.

Content

Abstract.....	i
Resumen	iii
Content.....	vi
Glossary of terms	viii
1. State of art.....	1
1.1 Medical context: Screening for disease	1
1.1.1 Historical background and current situation.....	1
1.1.2 Definitions and general principles of screening	4
1.1.3 Benefits and disadvantages of screening	7
1.1.4 Screening in ophthalmology.....	9
1.1.5 Screening in skin cancer	11
1.1.6 Image modalities for screening.....	12
1.2 The role of telemedicine in screening.....	15
1.3 Computer-aided detection and diagnosis in screening	18
1.3.1 Preprocessing and image enhancement	20
1.3.2 Segmentation	21
1.3.3 Feature extraction and classification	22
1.3.4 Convolutional neural networks.....	23
1.3.5 Performance of CAD systems	24
2. Objectives	26
2.1 Tele-screening tool	27
2.2 OCT enhancement algorithm for skin images.....	28
2.3 CAD system for glaucoma based on CNN.....	28
2.4 Context	29
3. Tele-screening tool	31
3.1 Introduction	32
3.2 Glaucoma diagnosis.....	33
3.3 Methods	38
3.4 Results and discussion	43
3.5 Conclusions	51
4. Enhancement for OCT skin imaging.....	52
4.1 Introduction	53
4.2 Method.....	55
4.2.1 Wavelet Compounding Adaptive Noise Filter (WCAN)	55
4.2.2 OCT Imaging System	57
4.2.3 Data.....	58
4.2.4 Quantitative Evaluation	59

4.2.5	Qualitative Assessment	61
4.3	Results and Discussion	64
4.3.1	Quantitative evaluation	64
4.3.2	Qualitative Evaluation	68
4.4	Conclusion	70
5.	CAD system for glaucoma disease	71
5.2	Materials and Methods	76
5.2.1	Study Datasets	76
5.2.2	Preprocessing.....	79
5.2.3	CNNs used and Transfer learning	80
5.2.4	Performance metrics	83
5.2.5	Training and testing processes.....	83
5.2.6	Implementation	84
5.3	Results and discussion	84
5.3.1	Hyper-parameter selection.....	84
5.3.2	CNN algorithm selection	84
5.3.3	Ten-fold cross validation	86
5.3.4	CNN/human evaluator performance.....	87
5.3.5	Datasets performance influence.....	90
5.3.6	Integration of medical data and CNN.....	91
5.4	Conclusions	93
6.	Contributions	94
	References.....	96
	Publications.....	124
Journal articles		124
Conference Proceedings		124
Software Registry		125
	Agradecimientos	126

Glossary of terms

AMD	Age-related macular degeneration
AUC	Area Under the (ROC) Curve
AK	Actinic Keratosis
BCC	Basal Cell Carcinoma
CAD or CADe	Computer-Aided Detection
CAD or CADx	Computer-Aided Diagnosis
CDR	Cup to disc ratio
CF	Color Fundus Images
CNN	Convolutional Neural Networks
CNR	Contrast to Noise Ratio
CT	Computed Tomography
DEJ	Dermis Epidermis Junction
DR	Diabetic retinopathy
EMR	Electronic Medical Record
ENL	Equivalent Number of Looks
FOM	Figure of Merit
HMF	Hybrid Median Filter
HRT	Heildeberg Retina Tomography
IOP	Intraocular pressure
ISNT	Inferior-Superior-Nasal-Temporal rule
MM	Malignant Melanoma
MVC	Model View Controller pattern design
NMSC	Non Melanoma Skin Cancer
NSC	National Screening Committee in the United Kingdom
OBNLM	Optimized Bayesian Nonlocal Means algorithm
ONH	Optic Nerve Head
OCT	Optical Coherence Tomography
RNFL	Retinal Nerve Fibre Layer
ROC	Receiver Operating Characteristics
SCC	Sqamous Cell Carcinoma
SNR	Signal to Noise Ratio
SS OCT	Swept-Souce Optical Coherence Tomography
WCAN	Wavelet Compounding Adaptive Noise filter
WHO	World Health Organization
WVMF	Wavelet Multi-Frame denoising algorithm

Chapter 1

State of art

1.1 Medical context: Screening for disease

1.1.1 Historical background and current situation

Screening for disease has been a widely accepted practice in healthcare since the twentieth century. In 1917 the Division of Psychology within the Medical Department of United States of America (USA) Army during World War I accomplished one of the oldest screening programs recorded. The purpose was to screen potential soldiers for psychiatric disorders before allowing them to join the US Army (Wells 1920). Later on, in 1944 the US Army also developed a theory for the conception and interpretation of screening tests. They proposed for the first time a quantitative index that allowed to assess the accuracy or performance of a screening program $TP/(TP+FN+FP)$ where TP=true positive, FN=false negative and FP=false positive (Morabia and Zhang 2004). Another remarkable program in that period was the diabetes screening done in New England (USA) in 1946-47 (Morabia and Zhang 2004; Wilkerson and Krall 1947) where 70.6% of the inhabitants in Oxford performed a urine and blood glucose test, resulting in a 1.7% prevalence of diabetes. In the USA in the late 50s a major concern with chronic disease

In the 1950s and early 1960s, the benefits of screening were demonstrated by the use of mass miniature radiography (MMR) for the identification of individuals with tuberculosis. In many western countries, especially in the USA and the United Kingdom (UK), the use of this technology became widespread after the introduction of effective treatments for this disease. For example in Liverpool (England) from 1953 to 1974, the Medical Office of Health set up a MMR campaign during which more than 80% of the population were screened (Holland and Stewart 2005). In 1957 an ambitious pulmonary tuberculosis campaign in Glasgow (Scotland) screened 76% of the adults, over 700,000 in total, where 2,200 were found active cases and provided treatment of more than 1,000 in hospital (Ian Levitt 2003). The mass screening of breast cancer are linked with the availability of mammography in the 1960s. After preliminary studies in December 1963, the Health Insurance Plan (HIP) of Greater New York launched a randomized clinical trial which lasted for almost 25 years and ended in 1986 including 62 000 women aged 40 to 64 years. The HIP study found out a 30% reduction in mortality from breast cancer during the first 10 years and follow up (Morabia and Zhang 2004). However, the benefits of breast cancer screening even today are controversial due to the overdiagnosis and the effect on the rate of death presented in several studies (Bleyer and Welch 2012; Marmot et al. 2013).

The Table 1.1 contains a summary of relevant historic screening programs in USA and UK.

Table 1.1. Review of relevant screening programs during the twentieth century in USA and UK.

Disease	Location	Years	Goal/Description (references)
Psychiatric disorders	USA	1917-18	Performed to exclude subjects with psychological disorders from young men eligible to join the United States army.((Wells 1920; Morabia and Zhang 2004))
Cervical cancer	New York City (USA)	Late 1940s	Some preliminary programs in late 1930s. After Papanicolaou and Traut first reported the usefulness of Papanicolaou (“Pap”) smear for detecting neoplastic cervical cells in 1943 it was widely used as the screening tool. (Morabia and Zhang 2004)
Diabetes	Oxford (Massachusetts, USA)	1946-47	Implemented to determine the prevalence of diabetes in a typical American community, evaluate methods for mass large screen diagnosis, inculcate the need of the periodic examinations for diabetes and start treatment by the family physician. ((Wilkerson and Krall 1947; Morabia and Zhang 2004))
Tuberculosis	Liverpool (England), Glasgow (Scotland)	1950s-60s	Special campaigns of mass radiography. Large reduction of cases thanks to the campaigns. (Holland and Stewart 2005)
Newborn diseases	USA	1960s	Executed to identify the presence of Phenylketonuria, a generic metabolic disorder. (Stiffler and Dever 2015)
Breast cancer	Greater New York (USA)	1963-1986	Health Insurance Plan (HIP) randomized clinical trail The screening examination consisted of a film mammography and an independently conducted physical examination of the breasts done by a physician. (Morabia and Zhang 2004)

In 1968 the Minister of Health in the UK created the Joint Standing Sub-Committee of Screening in Medical Care. Although with mainly advisor competences, the committee started to review the evidences of any screening program and make recommendations of the procedures to comply before a program could be introduced in the National Health Service. The sub-committee developed its activities until 1980. In 1996 and to identify areas of future development, the National Screening Committee (NSC) was founded. The NSC¹ represents a referent center for all considerations of screening in the UK and is responsible for introducing proven screening programs and for evaluating them. Nowadays the NSC has recommendations of 109 conditions which are reviewed every three years based on published evidence, consultation and stakeholder responses and the consideration of the committee members. There are currently 11 managed NHS (National Health Service) population screening programs in England. Only two other countries have similar organizations worldwide, the Netherlands and New Zealand. The Center for Population Screening² (CvB) in the Netherlands manages and coordinates the eight national population screening programs, and the National Screening Unit³³ (NSU) in New Zealand has currently six screening programs and one quality improvement program. The Table 1.2 details the current programs by location.

Table 1.2. Current Screening programs of countries with a National Reference Centre.

Location	Center	Screening Programs
England	NSC	Antenatal and newborn: sickle cell and thalassaemia, fetal anomaly, infectious diseases in pregnancy, newborn and infant physical examination, newborn blood spot and newborn hearing. Adult screening: diabetic eye, abdominal aortic aneurysm, breast cancer, cervical cancer and bowel cancer.
The Netherlands	CvB	Antenatal and newborn: Down's syndrome screening, the 20-week ultrasound, the heel prick program, hearing screening and pregnancy blood screen. Adults screening: cervical cancer, breast cancer, bowel cancer.
New Zealand	NSU	Antenatal and newborn: blood screen, Down's syndrome (optional), metabolic disorders (heel prick) and hearing. Adults: breast cancer, cervical cancer and bowel cancer.

¹https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/649986/Screening_in_the_UK_making_effective_recommendations_2016_to_2017.pdf

²The Dutch Centre for Population Screening Website.

https://www.rivm.nl/en/About_RIVM/Organisation/Centres/Centre_for_Population_Screening.

³ New Zealand National Screening Unit Website. <https://www.nsu.govt.nz>

The World Health Organization on behalf of the European Observatory on Health Systems and Policies published in 2006 a review of the screening practice in 28 countries (Holland, Stewart, and Masseria 2006). The report highlight that few countries have a single national organization to supervise screening practice and policy (like UK or the Netherlands) and that the screening practice can vary widely as a result of the regional or local health service provision in many countries of the EU. Regarding cancer screening, the Council of the EU issued recommendations setting out principles of best practice in the early detection of cancer¹. The Council recommendation called on all EU countries to take common action to implement national, population-based screening programs for breast, cervical and colorectal cancer. The last report supported by the European Commission² describes the status of the implementation of the screening programs (Table 1.3).

Table 1.3. Situation of screening programs that are ongoing, piloted or planned for breast, cervical and colorectal cancer in the EU in 2017 (International Agency for Research on Cancer, CPO Piemonte and University Hospital “Città della Salute e della Scienza,” and Mass Screening Registry/ Finnish Cancer Registry 2017).

Cancer	# EU members	Gender	Age group	% population in age group covered
Breast	25	Women	50-69	95
Cervical	22	Women	30-59	72
Colorectal	23	Men/Women	50-74	72

In Spain the 177 Commission of Public Health³ established in 2010 the basic criteria and requirements that could serve as a guide to the implementation of a screening program in all health services provided by the different Autonomous Communities of the country. Specific guides have also been established for screening in neonatal programs⁴ and cancer programs⁵.

1.1.2 Definitions and general principles of screening

Over the years, since its introduction, the concept of screening has been defined in different ways trying to focus on diverse aspects. One of the most accepted definitions was created by the United States (Commission on Chronic Illness 1957) that defined screening as “The presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment.” This statement

¹ <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32003H0878&from=EN>

² “Cancer Screening in the European Union (2017).” European Union Public Health Programme.

³ https://www.msssi.gob.es/profesionales/saludPublica/prevPromocion/docs/Cribado_poblacional.pdf

⁴ <https://www.msssi.gob.es/profesionales/saludPublica/prevPromocion/cribadoNeonatal.htm>

⁵ <https://www.msssi.gob.es/profesionales/saludPublica/prevPromocion/cribadoCancer.htm>

was also formally adopted by the WHO Regional Committee for Europe (Wilson and Jungner 1968). Other relevant definition was proposed by the UK NSC in its First Report¹ in 1998, “Screening is the systematic application of test or inquiry to identify individuals at sufficient risk of a specific disorder to warrant further investigation or direct preventive action among persons who have not sought medical attention on account of symptoms of that disorder”. In 2000, in the Second Report², the UK NSC changed some aspects of the definition, “Screening is a public health service in which members of a defined population, who do not necessarily perceive that they are at risk of, or are already affected by, a disease or its complications, are asked a question or offered a test to identify those individuals who are more likely to be helped than harmed by further tests or treatment to reduce the risk of disease or its complications”. We can see how the NSC puts the focus on the reduction of the risk of the disease and in the view of the screening as public health service. Apart from the definitions, another useful way to approach the different roles of screening is through the different types of screening (Table 1.4) that Wilson and Jungner described (Wilson and Jungner 1968).

Table 1.4. Types of screening (Wilson and Jungner 1968)

Type	Description
Selective	Screening of selective high-risk groups in the population.
Mass public	Large-scale screening of whole population groups.
Multiple (or multiphasic)	Application of two or more screening tests in combination to large groups of people.
Surveillance	Long-term process where screening examinations are repeated at intervals of time.
Case-finding	The main object is to detect and bring patients to treatment in contrast to epidemiological surveys where although screening tests can be made the aim is to find the prevalence or incidence of a disease.

Thanks to the experience of all previous programs and studies in the 1960s the first reviews with principles and procedures of screening were issued. In 1961 Thorner and Remein of the US Public Health Service published the first comprehensive review (Thorner and Remein 1961). Subsequently, in a report commissioned by the World Health Organization (WHO) J.M.G. Wilson and G. Jungner (Wilson and Jungner 1968) attempted to establish the screening criteria to guide the selection of principles (Table 1.5) that would be suitable for screening. The stage of the screening in relation with the condition of the development of the disease was also highlighted in the report. For instance, screening for the early diagnosis of malaria presents different problems with respect to the screening of a chronic disease like diabetes.

¹ “First Report of the National Screening Committee.” Health Departments of the United Kingdom.

<http://aogm.org.mo/assets/Uploads/aogm/Guidelines/NHS/National-Screening-Committee-UK-1998.pdf>

² “Second Report of the UK National Screening Committee.” Health Departments of the United Kingdom.
<http://aogm.org.mo/assets/Uploads/aogm/Guidelines/RCOG---UK/Second-report-of-UK-NSC.pdf>

Table 1.5. Wilson and Jungner classic principles of early diagnosis criteria. Based on (Wilson and Jungner 1968)

The condition sought should be an important health problem.
There should be an accepted treatment for patients with recognized disease.
Facilities for diagnosis and treatment should be available.
There should be a recognizable latent or early symptomatic stage.
There should be a suitable test or examination.
The test should be acceptable to the population.
The natural history of the condition, including development from latent to declared disease, should be adequately understood.
There should be an agreed policy on whom to treat as patients.
The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
Case-finding should be a continuing process and not a “once and for all” project.

Four decades later several authors (Andermann et al. 2008) revisited this public health guide to incorporate a synthesized view of the new criteria proposed since then (Table 1.6).

Table 1.6. Synthesis of screening for disease emerging criteria proposed by (Andermann et al. 2008)

The screening program should respond to a recognized need.
The objectives of screening should be defined at the outset.
There should be a defined target population.
There should be scientific evidence of screening program effectiveness.
The program should integrate education, testing, clinical services and program management.
There should be quality assurance, with mechanisms to minimize potential risks of screening.
The program should ensure informed choice, confidentiality and respect for autonomy.
The program should promote equity and access to screening for the entire target population
Program evaluation should be planned from the outset.
The overall benefits of screening should outweigh the harm.

The UK NSC in its Appendix C of the Second Report published the criteria “for appraising the viability, effectiveness and appropriateness of a screening program”. The criteria were based on the criteria proposed by WHO in 1968, taking into account also “the more rigorous standards of evidence required to improve effectiveness and greater concern about the adverse effect of healthcare”. In the following table we can see a summary of all the 19 criteria that should be met before screening for a condition according to the UK NSC.

Table 1.7. UK NSC criteria for appraising the viability, effectiveness and appropriateness of a screening program. (Health Departments of the United Kingdom 2000)

Area	Description
The condition	It should be an important health problem with an epidemiology adequately understood, with a disease marker and an early symptomatic stage. Besides, all cost-effective primary prevention should have been implemented as far as practicable.
The test	There should be simple, safe, precise, acceptable by population, with suitable cut-off level defined and validated. There should be an agreed policy for the further steps for positive test results.
The treatment	There should be an effective treatment and evidence of the better outcomes of its early application. Also agreed evidence based policies covering the individuals that should receive treatment and the clinical management of the condition and patient outcomes should optimized before the screening programme.
The screening programme	There should evidence of its effectiveness reducing mortality and morbidity, the outcome must be of value and well understood by the screened persons, the complete programme should be clinically, socially and ethically acceptable by all actors. In relation with the costs, the program should balance the opportunity cost in relation with the expenditure on medical care as a whole. The programme should be manage and monitor considering quality standards and should include all the staffing and facilities needed.

1.1.3 Benefits and disadvantages of screening

Screening has become more present in the health systems in recent years. The demand of these services are increasing sometimes due to political, public or commercial pressure. But not always this pressure is supported by scientific evidences and it is not proven the efficacy of the output of the screening in all the conditions. The Council of Europe in a recommendation to governments of member states on screening as a tool of preventive medicine highlights some of the disadvantages and adverse effects of screening (Table 1.8) and differentiates various types of problems concerned with ethical and legal issues, selection of diseases suitable for screening, economic aspects, quality assurance, organization of a screening program and scientific research.

Table 1.8. Adverse effects of screening ¹

Stigmatization and/or discrimination of (non) participants
Social pressure to participate in the screening and undergo the intended treatment/intervention;
Psychological distress where there is no cure for the disease or where the treatment and/or intervention is morally unacceptable to the individual concern
Exposure to physical and psychological risks with limited health gains
Creation of expectations which probably cannot be fulfilled
Individuals who are positively screened might experience difficulties such as access to insurance, employment, etc.;
Severe side effects of invasive clinical diagnosis of false positives
Delay in diagnosing false negative
Unfavorable cost-benefit relationship of a screening program.

The ethical issues are specially challenging in the increasing implantation of genetic screening. The psychological, behavior and perceived risk perspectives of these programs were reviewed in (Heshka et al. 2008), pointing out that genetic testing does not seem to lead to adverse consequences for carriers and non-carriers of the conditions studied (hereditary nonpolyposis colorectal carcinoma, hereditary breast and ovarian cancer, and Alzheimer disease). Other studies explore the ethical implications of introducing non-invasive prenatal diagnostic tests (de Jong et al. 2010) and its implications in selective abortion and the debate of the ‘right not to know’ of the future child.

The health professionals are becoming more sensitive to the possible risks of screening based on the experience of many years of programs and reviews. The situation could become problematic when the pressure comes from the screening industry and the media, like in the case of the whole body screening with computed tomography (CT) that was the focus of a major advertising campaign in recent years in the United States. (Swensen 2003) presented the example of “700 ancillary findings including four renal cell carcinomas, three breast cancers, two lymphomas, two gastric tumors, one pheochromocytoma, and 114 abdominal aortic aneurysms” using whole body screening with CT that at the end were most of them false positives. The need of randomized trials with cost-effectiveness outcomes by condition is at the end the only way to produce an answer about the merit of each screening. For example, in the case of lung cancer screening using low dose computed tomography, a systematic review of the evidence regarding the benefits and harms (Bach et al. 2012) concludes that may benefit individuals at an elevated risk for lung cancer (“smokers and former smokers who are age 55 to 74 and who have smoked for 30 pack years or more and either continue to smoke or have quit within the past 15 year”), but uncertainty exists about

¹ “Council of Europe (1994) Recommendation No R(94)”

<https://wcd.coe.int/com.instranet.InstraServlet?command=com.instranet.CmdBlobGet&InstranetImage=534532&SecMode=1&DocId=514336&Usage=2>

potential harms and the generalizability of results. Similarly, (Raffle et al. 2003) showed that for the cervical screening program in Bristol around 1000 women needed to be screened for 35 years to prevent one death.

1.1.4 Screening in ophthalmology.

Screening for retinal diseases, in particular diabetic retinopathy (DR) and age-related macular degeneration (AMD), is one of the most common and successful application of early diagnosis of eye diseases until now. DR is an eye disorder that affects the retinal vasculature and gradually develops in patients with diabetes. Visual problems derived from DR provoke dark spots in the vision associated with the areas of the retina affected. Globally, the number of people with DR will grow from 126.6 million in 2010 to 191.0 million in 2030 (Zheng, He, and Congdon 2012). DR is the leading cause of vision loss in middle-aged and elderly people globally (Ting, Cheung, and Wong 2016). According to WHO it is estimated that 4.8% of the number of cases of blindness worldwide are caused by DR (Resnikoff et al. 2004). Among the people affected by diabetes the cumulative incidence of DR increase with time; for the population with 25 years suffering Type 1 diabetes the rate reach 97% (Klein et al. 2008). Due to the increasing number of people with diabetes worldwide (estimated in 415 million in 2015 and in 642 million in 2040 (Ogurtsova et al. 2017)) the prevalence rate of DR is estimated to increase rapidly in the near future (Saaddine et al. 2008). AMD accounts for 8.7% of all blindness worldwide and is the most common cause of blindness in developed countries especially in people older than 60 years old (W. L. Wong et al. 2014). AMD affects the central vision preventing patients to perform detailed activities. The project number of people with AMD in 2020 is 196 million increasing to 288 million in 2040 (W. L. Wong et al. 2014).

DR or AMD are generally asymptomatic at early stages and as the disease advanced irreversible damage to the retina occurred (Ciulla, Amador, and Zinman 2003; H. R. Coleman et al. 2008). Early detection of DR in individuals with diabetes is considered critical in preventing visual loss. Several studies have confirmed that the clinical outcome is better if patients with diabetes are screened and treated early, saving years of vision and reducing societal costs (Ciulla, Amador, and Zinman 2003; Mohamed, Gillies, and Wong 2007). AMD screening programs are receiving great interest in the last years with the first implementations showing its effectiveness in reducing the number of patients with blindness although presenting discrepancies in the cost-effectiveness of these methods (Bats et al. 2014; Chan et al. 2015; Tamura et al. 2015).

Other promising condition for screening is glaucoma. Glaucoma is a group of optic neuropathies characterized by progressive degeneration of retinal ganglion cells (Weinreb, Aung, and Medeiros 2014). Glaucoma is the leading cause of global irreversible vision loss with a prevalence for population aged 40-80 estimated in 3-4% (Tham et al. 2014; Burr et al. 2007). The number of people with glaucoma worldwide was estimated in 64.3 million in 2013, increasing to 76.0 million in 2020 and 111.8 million in 2040 (Tham et al. 2014).

Because glaucoma is an asymptomatic condition until a relatively late stage the diagnosis is frequently delayed. Population-level surveys suggest that only 10-50% of people with glaucoma are aware they suffer the disease (Weinreb, Aung, and Medeiros 2014). As early diagnosis of the condition and treatment can prevent vision loss, glaucoma screening has been tested in numerous studies worldwide (Burr et al. 2007; Tielsch et al. 1991; C. Fleming et al. 2005; Maul and Jampel 2010; D. Zhao et al. 2017). Current studies show that glaucoma screening can be cost-effective in risk population (family history, black ethnicity, age) and can be improved using a test with initial automated classification followed by the assessment of a specialist (Burr et al. 2007). The next figure presents a representation of a scene viewed by a person with some of these conditions.

Figure 1.1. A scene as it might be viewed by a person with (a) normal vision (b) diabetic retinopathy (c) Age-related macular degeneration and (d) glaucoma. Source: images courtesy of NIH, National Eye Institute.

An example of the method followed in a real glaucoma screening program could be the Salzburg-Moorfields Collaborative Glaucoma study that was developed over a period of eight years (Hitzl et al. 2006).

- 4864 individuals were screened.
- The inclusion criteria were age ≥ 40 , best spectacle correction visual acuity $>6/9$, refraction from -6.00 to $+4.00$ diopters and difference in refraction between both eyes <3.00 diopters.
- Exclusion criteria considered several eye conditions like pseudophakia, glaucoma therapy among others.
- The examinations included a complete eye examination including refraction, best-corrected visual acuity, applanation tonometry, biomicroscopy of the anterior segment,

fundus examinations with assessment of the optic nerve head and the subjective cup-to-disc ratio, history of existing eye diseases and current topical treatment. Furthermore the visual Field Test Analyzer.

- They used the diagnostic criteria suggested by European Glaucoma Society.

Other conditions that have been included in screening studies are retinopathy of prematurity (J. Chen et al. 2011; Haines et al. 2005), Alzheimer's disease using retinal photographic analysis (Frost et al. 2013) or visual and ocular disorders in school children (Baltussen, Naus, and Limburg 2009; Kvarnström, Jakobsson, and Lennerstrand 2001; He et al. 2007).

1.1.5 Screening in skin cancer

Malignant Melanoma (MM) and non-melanoma (Bowen's disease, squamous cell carcinoma and basal cell carcinoma) skin cancer (NMSC) are the most common types of cancer in the Caucasian populations (Diepgen and Mahler 2002; Apalla et al. 2017). The incidence of MM is increasing in most European countries in generations born up to the end of the 1940s, with a stabilization or decline in other countries like Australia, New Zealand, USA, Canada and Norway (Erdmann et al. 2013). WHO¹ estimates currently between 2 and 3 million NMSC and 132,000 MM, which means that one in every three cancers diagnosed is skin cancer. In the last years several campaigns studied the benefits of early detection of skin cancer (De Rooij et al. 1996; Wolff 2009). Like in other conditions current analyses are not conclusive about the reduction of skin cancer mortality rate or to assess the balance of benefits and harms of the skin cancer screening (Stang and Jöckel 2016; Bibbins-Domingo et al. 2016). It is important to highlight that biopsy, and the correspondent histopathology result, is considered the reference standard to diagnose skin cancer, but when the lesion is considered benign the clinical diagnosis of the dermatologist, without biopsy, is accepted.

As an example of a skin cancer screening we can review the skin cancer screening in Northern Germany from July 2003 to June 2004 (Waldmann et al. 2012):

- 12 month skin cancer screening intervention.
- The study consisted on a whole-body examination conducted by either general practitioners and/or dermatologists
- Eligible persons were residents older than 20 years and policyholders of statutory health insurance and not receiving skin cancer care.
- If a non-dermatologist found a suspicious skin lesion, the patient was referred to a dermatologist to receive a second whole-body examinations.
- Participation was voluntary and the participants gave written informed consent for data storage and analysis.

¹ <http://www.who.int/uv/faq/skincancer/en/>

1.1.6 Image modalities for screening

Many current screening programs use medical images as an indispensable diagnostic tool. The great advance of medical imaging and image processing, their intensive use in the majority of the specialties to aid in the detection and follow-up of most conditions and their availability in the health provider centers have driven medical imaging to be one of the tests typically performed in a screening campaign. In the next table we can see a review of image modalities used in different types of screening programs.

Table 1.9. Image modalities used in screening. CT, computed tomography. US, ultrasound. MRI, Magnetic resonance imaging. CF, Color fundus photography. HFUS, High Frequency Ultrasound. OCT, Optical Coherence Tomography. RCM, Reflectance Confocal Microscopy.

Screening	Image modalities
Breast cancer	Mammography radiography, US, MRI (Smith et al. 2015; Berg et al. 2012)
Colorectal cancer	Computed Tomography colonography (Smith et al. 2015)
Lung cancer	Low-dose helical CT, Chest radiography (Smith et al. 2015; “The National Lung Screening Trial: Overview and Study Design” 2011)
Fetal anomaly	US (Salomon et al. 2011)
Diabetic Retinopathy, AMD	CF (Garg and Davis 2009; Bats et al. 2014)
Melanoma	Dermoscopy (Carli et al. 2004; Iyatomi et al. 2008)
NMSC	HFUS, OCT, RCM (Ulrich et al. 2007)

Among all previous modalities, in this PhD thesis proposal, we will focus on image processing methods for Color Fundus (CF) and Optical Coherence Tomography (OCT) images. While CF is basically used in retinal imaging, OCT can be used in retinal and skin imaging (among many other fields).

In CF photography, a color image of the interior surface of the eye is acquired and recorded using a fundus camera. A CF camera encapsulates a set of special lenses attached to a flash digital camera. The optics of a fundus camera are similar to those of an indirect ophthalmoscope. A CF camera is usually described by the angle of view, the optical angle accepted by the lens. The angle of view could vary from 20° or less in narrow angle fundus camera, to 45° until 140° in wide angle fundus cameras. CF allows a detailed view of retina, the retina vasculature, and the optic disc that can help to detect eye diseases even in asymptomatic early stages. CF is the most widely used technique to image the retina and has been often used in combination of image processing algorithms to aid in the detection and grading of eye diseases (M. Niemeijer et al. 2010; Bock et al. 2010; Burlina et al. 2017). The next figure shows a typical fundus photograph and a non-mydratic Topcon retinal camera.

Figure 1.2. (a) Fundus photograph of a healthy right eye of a 25-year old male (By Mikael Häggström, used with permission, Source: DOI:10.15347/wjm/2014.008) (b) A non-mydratic Topcon retinal camera (By Jason Ruck, Source: [http:// https://commons.wikimedia.org/wiki/File:Retinal_camera.jpg](http://https://commons.wikimedia.org/wiki/File:Retinal_camera.jpg))

Optical coherence tomography (OCT) (D. Huang et al. 1991) is an optical imaging technique that allows cross-sectional views of highly scattering biological in vivo tissue in real time with micrometer resolution and depth up to several millimeters. OCT was created with the goal to enable noninvasive optical biopsy, allowing the in situ imaging of tissue microstructure with a resolution similar to histology and without the need for tissue excision and processing (Drexler et al. 2014; Fujimoto et al. 1995). OCT systems are based on the Michelson Interferometer. The light from a light source is split by a beam splitter into the sample and the reference arm, the interference of the backreflected light from the two beams is captured by a detector and processed. In the figure 1.3 we can see the two modalities of OCT developed until now, Time Domain (TD) OCT and Fourier Domain (FD) OCT. TD OCT was the first implementation of the technique (D. Huang et al. 1991). In this modality a low coherence broad-band light source is used to allow the creation of a detectable interference signal when the optical path length of the sample and the reference arm is matched to the narrow coherence length. The interference signal at different depths and recorded at different time delays is demodulated to produce a reflectivity depth profile (A-scan). The next step in this technology was FD OCT that uses a broad-bandwidth light source and detects the interference spectrum from the interferometer, using a spectrometer and a line scan camera (Fercher et al. 1995; Drexler and Fujimoto 2008). Finally the swept source FD OCT uses an interferometer with narrow band-width light source, which is frequency swept in time. The variation in frequency with time essentially labels different time delays in the light beam, which can be detected by interference and Fourier transforming (Chinn, Swanson, and Fujimoto 1997; Drexler and Fujimoto 2008).

Figure 1.3. Different OCT modalities. OCT can be classified in Time Domain (TD) and Fourier Domain (FD) systems. OCT FD can be spectrometer based or swept source based. Scheme from (Drexler et al. 2014)

OCT is broadly used in multiple biomedical and clinical fields, such as cardiology as a catheter-based imaging system for the guiding in coronary interventions (Bezerra et al. 2009), to monitor the growth and development of biological tissues over time (Boppart et al. 1996; Men et al. 2016), to detect dysplasia and occult malignancy in Barrett's esophagus (X. D. Li et al. 2000; Cobb et al. 2010), to aid in the diagnosis and treatment of multiple of laryngeal disorders (Burns 2012; Armstrong et al. 2006), as an adjunct to bronchoscopy in the diagnosis of lung cancer (Michel et al. 2010), among others.

In ophthalmology OCT has become a clinical standard because the eye is easily imaged and there is no other noninvasive image system with this resolution. OCT can aid in the identification of disease in early stages before physical symptoms and irreversible vision loss occurs. Besides, repeated imaging can be performed to allow the evaluation of the disease progression or monitor the effectiveness of therapy. The first *in vivo* retinal imaging were obtained in 1993 (Swanson et al. 1993; Fercher et al. 1993). The layers of the retina and other architectural features can be visualized with higher resolution than even before. OCT has been applied to monitor a great variety of eye conditions of macular diseases (Puliafito et al. 1995; Srinivasan et al. 2006) including macular edema (Diabetic Retinopathy Clinical Research Network et al. 2007), macular hole (Hee et al. 1995; Philippakis et al. 2018) or age-related macular degeneration (Manjunath et al. 2011). The retinal nerve fiber layer thickness is an indicator of glaucoma that can be assessed and quantified in normal and glaucomatous eyes with OCT and correlated with conventional measurements of the optic nerve structure and function (Schuman et al. 1995; Medeiros et al. 2005; Leung et al. 2010).

In dermatology OCT is a useful tool that has been used to study a variety of dermatological disorders like systemic lupus erythematosus (T Gambichler et al. 2007), bullous diseases (Mogensen et al. 2008), psoriasis and contact dermatitis (Morsy et al. 2010; Welzel, Bruhns, and Wolff 2003), wound healing (Greaves et al. 2014) or vascular skin lesions (Salvini et al. 2008). One of the most prominent application of OCT in dermatology is the assessment of non-melanocytic epithelial tumors: actinic keratosis (AK) (Mogensen et al. 2009; Jennifer K. Barton et al. 2003), basal cell carcinoma (BCC) (Mogensen et al. 2009; Gambichler, Orlikov, et al. 2007; Hussain, Themstrup, and Jemec 2015; Olmedo et al. 2006) and squamous cell carcinoma (SCC) (Hatta et al. 2010).

One of the most active research field in OCT technology nowadays is the optical coherence tomography angiography (OCTA), a new non-invasive imaging technique that allows the generation of angiographic volumetric images in few seconds. This new technique is showing potential applications in retinal vascular diseases (de Carlo et al. 2015). Another relevant area, that links with this project, is related with several recent promising studies that show the potential role that low-cost handheld OCT systems could have in the screening of diseases in early stages (Shelton et al. 2014; Jung et al. 2011).

1.2 The role of telemedicine in screening

The advances in information and communication technology (ICT) in the last century created new opportunities to deliver health care services remotely. Since the introduction of the term in the 1970s there have been many attempts to define telemedicine. WHO adopted the following description: “The delivery of health care services, where distance is a critical factor, by all health care professional using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of health care providers, all in the interests of advancing the health of individuals and their communities” (WHO 2010). WHO also identifies four key elements when we refer to telemedicine: (1) Its purpose is to provide clinical support. (2) It is intended to overcome geographical barriers, connecting users who are not in the same physical location. (3) It involves the use of various types of ICT. (4) Its goal is to improve health outcome. According to the last survey of WHO globally the telemedicine services more established are teleradiology, telepathology, teledermatology and telepsychiatry. There are other health fields with also relevant initiatives in telemedicine but with less development: cardiology, ultrasonography, mammography, surgery, ophthalmology or diabetes among others. Telemedicine is usually divided in two basic types, store-and-forward and real time telemedicine. Store-and-forward telemedicine implies that the data related with the health service is stored previously to their future use by the expert or the health professional involved in the exchange. Real time telemedicine assumes that the participants of the health service are present at the same time, like in a videoconference teleconsultation. Most of the studies related with screening use store-and-forward telemedicine.

The telemedicine and e-health technology have played a significant role assisting in the provision of screening services. In all the most established screening services like in breast cancer (Lou et al. 1997; Pruthi et al. 2013), cervical cancer (Quinley et al. 2011) or lung cancer screening (Takizawa et al. 2001) there have been remarkable telemedicine initiatives.

The intensive use of imaging for the diagnosis and monitoring of eye diseases make ophthalmology an ideal field for telemedicine. Since its introduction in 1987 as part of a project to monitor retinal vessels during a space flight of the Space Shuttle Columbia (Hunter 1994) the teleophthalmology has been expanding in different areas like allowing the access to eye specialist in remote areas (Kumar et al. 2006; Bai et al. 2007), research and clinical trials (B. Li et al. 2015), distant learning projects (Stahl et al. 2009; Goldschmidt and Goodrich 2004) and, related with this project, in eye disease screening. The most common application until now has been in diabetic retinopathy screening (Liesenfeld et al. 2000; Neubauer et al. 2003; Raman et al. 2007; Kanagasingam Yogesan et al. 2000). According to (Das et al. 2015) DR is an ideal ophthalmic disease for telescreening and decision-making because it fits to Wilson and Jungner's criteria of screening for chronic diseases and the American Telehealth Association's screening categories. The programs normally start as part of a regular diabetic follow-up examination with the acquisition of fundus images obtained with a nonmydriatic digital retinal camera. The digital retinal images are sent securely through Internet to a specialist who performs remotely a report with the diagnostic of the status of the retinopathy. The report is finally sent to the patient through its primary care center or similar health center. The American Telemedicine Association (ATA) in its guide "*Telehealth Practice Recommendations for Diabetic Retinography*" (H. K. Li et al. 2011) created a series of recommendations for the design and implementation of teleophthalmology diabetic retinopathy programs. The previous guide divided a telescreening DR program in four broad elements to be care of: image acquisition, image review and evaluation, patient care supervision and image and data storage. Glaucoma is another eye condition that shows a promising potential for telescreening. A recent review of the subject (Thomas et al. 2014) including 45 studies which evaluate the use of telemedicine in screening concludes that teleglaucoma detects more cases of glaucoma than in-person examination and that it is an effective screening tool for glaucoma specially for remote communities. In the next figure there is list of all the different components usually present in teleglaucoma screening. We can find also telescreening studies in retinopathy of prematurity (Lorenz et al. 2009; Quinn et al. 2014; Richter et al. 2009) and in age-related macular degeneration (Bats et al. 2014; Duchin et al. 2015).

Figure 1.4. Standardized teleglaucoma equipment (Source: (Thomas et al. 2014) <https://doi.org/10.1371/journal.pone.0113779.t001>)

Components	Requirements
Human Resources	Staff: graders, Ophthalmic technicians, nurses, optometrist, physicians, glaucoma specialists/ophthalmologists
Information Technology	Secure Diagnostic Imaging (SDI) system
	Videoconferencing equipment
	Computer systems and software
	ISDN installation
Screening Equipment	Retinal camera
	Tonometer
	Devices to measure central corneal thickness
	Frequency Doubling Technology (FDT) or Humphrey Visual Field test
	Optical Coherence Tomography
	Slit lamp
	Gonioscope
	Retinal camera
	Tonometer
	Devices to measure central corneal thickness
Examinations	Medical & family history
	Visual acuity
	IOP
	CCT
	Pupil equal and reactive to light (PERL) or relative afferent pupillary defect (RAPD)
	Slit lamp
	Gonioscopy
	Visual field
	Fundus photographs
	OCT
	Ancillary tests

doi:10.1371/journal.pone.0113779.t001

Teledermatology is one of the most developed fields in telemedicine. As in ophthalmology, imaging has a major role in the skin examination, which partly explains its broad adoption. A throughout review of the use of teledermatology and teledermatoscopy (Warshaw et al. 2011) concludes that the diagnostic accuracy of clinic dermatology is still better than teledermatology, but that dermatoscopy-trained teledermatologist may be beneficial for isolated skin lesions. The skin conditions that the studies reported include rashes and circumscribed lesions (Edison et al. 2008; Bowns et al. 2006), pigmented skin lesions and nonpigmented skin lesion (Warshaw et al. 2009). Recently, the improvement in the capabilities in the mobile technology and the massive adoption of smartphones worldwide has expanded the adoption of these solutions for dermatology care. (Brewer et al. 2013) identified 229 dermatology-related apps in the most important app stores. The main categories were general dermatology references, self-surveillance/diagnosis, disease guide, educational aid and others.

One of the most prominent uses of teledermatology is in the early detection of skin cancer. One recent review that evaluated 21 studies (Finnane et al. 2017) stated that although robust implementation are needed, teledermatology services reduced waiting times to assessment and diagnosis and that patient satisfaction was high. The assessment of the recent studies shows that the accuracy of face to face dermatology consultations is still in general higher

than teledermatology, although the study remarks the methodological limitations of the studies, which make difficult to conclude if the differences in the diagnosis are due to the use of the technology itself or the expertise or training of the specialist involved in the studies. In screening, numerous studies have explored the use of this technology (Baumeister et al. 2009; Mahendran, Goodfield, and Sheehan-Dare 2005; Lowitt et al. 1998) . As in teledermatology in the last years the use of mobile phones has been extensively proposed for skin cancer screening studies (Lamel et al. 2012). A recent survey of 39 apps for melanoma skin cancer detection showed that only four of them provided tools to assess a skin lesion for likelihood of melanoma (Kassianos et al. 2015). The next step, as we will discuss later, seems to overcome these limitations and comes with the adoption of Deep Learning technology that in the future, outfitted in smartphones, could provide dermatologist-level classification of skin cancer detection (Esteva et al. 2017).

Teledermatology can be applied using clinical information and images or in combination with dermatoscopy digital images. Dermatoscopy (also known as dermoscopy or epiluminescence microscopy) is an imaging technique which allows physicians to visualize morphologic features that are not visible by naked-eye examination (Zalaudek et al. 2006). It is a noninvasive and non-expensive image modality that provides high reliability compared to face-to-face examinations has been used remotely either for melanocytic and non-melanocytic skin lesions (Şenel, Baba, and Durdu 2013; Piccolo et al. 2000). OCT is an image modality expanding in dermatology, as we described before, it is a valuable aid in the detection and monitoring of skin lesions but its high prize compared to dermoscopy, limits their use in screening and early detection. The use of portable handheld low-cost probes can improve the adoption of this technology and increase their use in primary care as has been already presented in recent studies (Shelton et al. 2014; Pande et al. 2017).

1.3 Computer-aided detection and diagnosis in screening

The term Computer-Aided detection (CAD or CADe) and diagnosis (CAD or CADx) refers to the technology that assists physicians in the interpretation of medical images for the detection and diagnosis of different pathologies. Although the first studies that reported quantitative analysis of medical images in the 1960s (Meyers et al. 1964; Lodwick et al. 1963) were not successful it was generally accepted that computers could improve the performance of specialists identifying abnormalities in images. The high expectations of this technology were not fulfilled at that time, but in the 1980s a new approach emerged. Computers could provide useful information to the physicians to aid in the diagnosis, but not to replace them. This new idea named computer-aid diagnosis was rapidly and widely accepted. The output of the CAD systems could provide even lower performance than physicians and still be useful if it offers complementary information that assists them in a better diagnosis. Nowadays CAD systems are a major research area and have been successfully used in screening programs of many different fields like breast cancer (Freer and Ullissey 2001; Rangayyan, Ayres, and Leo Desautels 2007; Nishikawa 2007), lung cancer

(Murphy et al. 2009; Y. Zhao et al. 2012; El-Baz et al. 2013), tuberculosis (Melendez et al. 2016; Muyoyeta et al. 2014), diabetic retinopathy (Mookiah, Acharya, Chua, Lim, et al. 2013; Sánchez et al. 2011) or skin lesions (W.-Y. Chang et al. 2013; Iyatomi et al. 2008; Masood and Ali Al-Jumaily 2013).

The main components of a CAD systems are: (1) preprocessing and image enhancement to improve the quality of the images to be processed (2) segmentation of relevant structures to limit the classification and focus on the region of interest of the condition to be assessed (3) feature identification to abstract findings in the images than can correlate with the presence or absence of a condition (4) and classification of the images through the features extracted. In the next sub-sections we will explore in more detail each component. The next figure shows an example of the different steps in a typical CAD system of DR diagnosis.

Figure 1.5. CAD system for DR diagnosis. The off-line system is performed only once and trains the system with features extracted from previously pre-processed and segmented images. For the training a manually classification of images (gold truth) is needed. The on-line system uses new images to evaluate and use the previous training parameters to classify them as normal or pathological. (Source: Mookiah et.al. (Mookiah, Acharya, Chua, Lim, et al. 2013))

1.3.1 Preprocessing and image enhancement

The first step in a CAD system is usually a preprocessing step where the images are prepared to improve the efficiency of the next stages and to reduce the variability between different data sources. During the acquisition of images it is common to use different devices at different locations and sometimes even different settings that can affect to the field of view, the resolution or the quality of the images. CAD systems are based on the training of massive datasets of images to learn the features that lately will detect in new images, images with poor quality can negatively affect in the training and in the posterior classification. Also, in early detection screenings based on the manually assessment of images by specialists the enhancement of the quality of the images represents a relevant step that can aid in the evaluation and diagnosis of the study.

In CAD systems with fundus retinal imaging techniques like histogram equalization (Sopharak et al. 2008), color remapping (Marrugo and Millán 2011) or contrast enhancement (Rasta et al. 2015) have been used to pre-process the images. But when the quality is too low sometimes is better strategy to discard the image because, even with the application of enhancement techniques, an accurate classification is not possible. There are several automatic image quality assessment implementations described in (Alan D. Fleming et al. 2006; Paulus et al. 2010).

One major research area in image enhancement in OCT is speckle noise reduction. Due to the use of spatially coherent illumination OCT images are affected by speckle phenomena that has a dual role as source of noise and carrier of information (Schmitt, Xiang, and Yung 1999; Bashkansky and Reintjes 2000). The speckle carrying signal information is the result of the back scattering of the incident photons while the speckle noise is caused by the random interference between multiple reflected photons coming from multiple directions. Speckle noise creates a grainy appearance in the OCT images that degrades the signal-to-noise (SNR) and limits its interpretation. Their properties are affected by the scale representation, the optical setting and the scattering properties of the biological tissue (P. Lee, Gao, and Zhang 2011). As source of information the speckle pattern statistics have been used to detect blood flow (Mariampillai et al. 2008; Srinivasan et al. 2010), differentiate living and non-living tissue (Gossage et al. 2006) or in optical elastography (Kirkpatrick, Wang, and Duncan 2006).

Speckle denoising is an active research field broadly developed during the last years. The different methods can be divided in those that modify the image configuration, where the optical setting or the scanning protocol can be adjusted, and those based on post-processing digital algorithms. In the first group we can find different forms of compounding techniques using the frequency (Pircher et al. 2003a), the angle of the incident light source (Iftimia, Bouma, and Tearney 2003; Desjardins et al. 2006) or multiple A-lines collected in a controlled way (Szkulmowski et al. 2012). Recent studies propose methods that combine image processing and modifications in the optical configuration to increase SNR even further

(Duan et al. 2014; W. Cheng et al. 2017). To compensate the effect of possible shifts in case of *in vivo* imaging weighted average of individual A-scans (Sander et al. 2005) or a combination of high-speed CMOS line scan camera and a cross correlation of different B-scans to a fixed reference one (Götzinger et al. 2011) has been proposed. One of the latest works suggests that the SNR improvement of pure angular compounding techniques are limited by optical aberrations (Winetraub et al. 2017).

In this work we developed a new technique for speckle reduction in OCT to improve the visualization of OCT skin images and their posterior interpretation for early detection of skin lesions.

1.3.2 Segmentation

The next step in a CAD system is the segmentation of relevant structures or regions necessary for the posterior extraction of features. In this study we focused on fundus images for the detection of eye conditions. We can divide the segmentation methods in (1) the localization and segmentation of retinal structures like optic disk, very useful in the detection of glaucoma (Joshi, Sivaswamy, and Krishnadas 2011; Sanchez et al. 2004) the fovea or the vasculature (Soares et al. 2006; Meindert Niemeijer, Abràmoff, and Ginneken 2009; Tan et al. 2017); (2) segmentation of retinal abnormalities like microaneurysms (earliest visible sign of retinal damage), hemorrhages (due to leakage of weak capillaries) (A. D. Fleming et al. 2006; Quellec et al. 2008), exudates, drusen among others (Abramoff, Garvin, and Sonka 2010; Meindert Niemeijer et al. 2007). In the next figure we present some examples of color fundus images of healthy and DR cases.

Figure 1.6. Typical healthy and different stages of DR fundus images: (a) Normal; (b) Mild Non-Proliferative DR (c) Moderate Non-Proliferative DR; (d) Severe Non-Proliferative DR; (e) Prolific DR; (f) Macular edema. (Source (Mookiah, Acharya, Chua, Lim, et al. 2013))

1.3.3 Feature extraction and classification

The feature extraction and their posterior use in the training of a statistical classifier are other very common components in current CAD systems. Features extracted from training examples images should be provided in order to construct the classifier. The classification can be divided in supervised classification, where a reference label is provided for its features, or unsupervised where there is not such a reference and the system must infer them in the training phase. Features can encode information at different levels of abstraction from pixel to patient level. In the case of pixel feature classification, it is usually performed using a supervised approach. During the implementation of a pixel feature classifier we need two different phases: (1) training phase where the classifier learns from labeled or well-known previous classifications; (2) testing or classification phase where the system classifies new unknown images (Abramoff, Garvin, and Sonka 2010). To detect lesions one approach used after a pixel classification, is performing a clustering of pixels into highly probable lesion regions and used different features to differentiate clusters (such surface area, length of major axis, standard deviation ...) (Meindert Niemeijer et al. 2007).

Automatic identification and quantification of retinal lesions algorithms based on previous approaches can be found in the main findings present in AMD, DR or glaucoma. In AMD, drusen is the most remarkable lesion and has been studied in several works (Mark J. J. P. van Grinsven et al. 2013; Liang et al. 2010). For DR the detection of red lesions (microaneurysms and hemorrhages) and bright lesions (exudates and cotton wool spots) have been extensively developed in numerous studies (Sopharak et al. 2008; Zhang et al. 2009; Larsen et al. 2003; Mane, Kawadiwale, and Jadhav 2015; García et al. 2010). In glaucoma the hallmark is cupping of the optic nerve head. The ratio of the optic disc cup and neuroretinal rim surface areas in fundus color images, called cup-to-disc ratio, is an important structural indicator for assessing the presence and progression of glaucoma (Abramoff, Garvin, and Sonka 2010). The assessment of the Optic Nerve Head (ONH) has been implemented by several groups (M. D. Abramoff et al. 2007; D. W. K. Wong et al. 2008).

The final output of CAD systems is a score of the likelihood of the presence of a disease based on the detection of previous lesions or the analysis of the complete image. In most developed systems a severity grading is also provided. Different solutions have been developed and tested in DR. In (Philip et al. 2007) a CAD system was tested with 6722 patients grading “disease/no disease” with 90.5% of sensitivity (86,5% in manual grading) and 67.4% specificity (95.3% in manual grading). (Alan D. Fleming et al. 2010) presented the evaluation of an automated grading software for DR with 33535 patients, detecting 100% of patients with proliferative retinopathy (180) and observable background retinopathy (324), 97.4% (1099/1130) with referable maculopathy and 99.2 % (1824/1827) with observable maculopathy. AMD automatic detection systems are in a more incipient state. (Liang et al. 2010; Agurto et al. 2011; Mookiah et al. 2014) described different implementations but with a small number of patients. In glaucoma the situation is similar. (Bock et al. 2010; Nayak et

al. 2009) created automatic systems for detecting glaucoma but were tested with small datasets.

1.3.4 Convolutional neural networks

Convolutional networks, commonly known as one of the most popular deep learning algorithms for image analysis, have become very rapidly a very successful alternative for analyzing medical images. These methods could be considered as the evolution of the supervised techniques, mentioned before, started at the end of the 1990s, where a training data sets of previously classified images are used to develop the system. Before deep learning revolutionized computer aid diagnosis systems the standard approach was based on feature extraction of relevant characteristics manually identified in the images and the implementation of a statistical classifier trained in a supervised way. The new deep learning paradigm implies that the computers can perform the feature learning and classification simultaneously. We can find usually in a deep learning algorithm a network (model) formed by many layers that transform an input data (images normally) to outputs (e.g. pathology present/absent). The most successful type of models for medical image analysis is a sub-class of neural networks called convolutional neural networks (CNN) that was introduced in the 1980s (Fukushima and Miyake 1982). The first relevant real-world application was in handwritten digit recognition with the network LeNet (Lecun et al. 1998). The recent advancements in core computing systems caused the exponential growth of the use of this technology first in computer vision where AlexNet (Krizhevsky, Sutskever, and Hinton 2012), the proposed CNN that won the challenge ImageNet in 2012, is considered the key breakthrough in this field. It is also relevant to highlight the role in the widespread use of deep learning of the availability of GPUs (Graphics processing unit), GPU-computing libraries (like CUDA or OpenCL) and several open source packages, like Caffe (Jia et al. 2014), Tensorflow (Abadi et al. 2016) or Theano (Bastien et al. 2012) that provide efficient implementations of very common CNN operations.

(Litjens et al. 2017) provided a thorough review of the current use of these techniques in medical analysis. The study mentions state-of-art applications of deep learning technology in the main topics of biomedical image processing: classification, object detection, segmentation or registration among others. The CAD systems and the screening programs have also progressively adopted the new approach. Several studies outperform state-of-art CAD systems on the mediastinal lymph node detection and interstitial lung classification (Shin et al. 2016), mammographic lesions (Kooi et al. 2017), lung nodules (J.-Z. Cheng et al. 2016) or skin lesions (Esteva et al. 2017).

CNN have been successfully applied to color fundus images achieving state-of-art performance or outperforming previous implementations. We can find since 2015 numerous studies applying CNN in retinal vessel segmentation (Huazhu Fu et al. 2016), image quality assessment (Mahapatra et al. 2016), segmentation of the optic disc and the optic disc cup (Zilly, Buhmann, and Mahapatra 2017), glaucoma detection (X. Chen et al. 2015), diabetic

retinopathy detection (Michael David Abràmoff et al. 2016), age-macular degeneration detection (Burlina et al. 2017) or hemorrhage detection (M. J. J. P. van Grinsven et al. 2016) among others.

1.3.5 Performance of CAD systems

To validate the CAD systems, measure its performance and compare it with human specialists several ratios have been proposed. The data set where the output ratios of the CAD systems are calculated must be kept separated from the data set used during the training of the algorithm.

The receiver operating characteristics (ROC) graph is a technique for visualizing, organizing and selecting classifiers based on their performance (Fawcett 2006). In a simple classification problem with two classes we can consider a class with score 0 as normal or negative and with score 1 as being abnormal or positive. In CAD systems the output is usually a score with the likelihood for each group detectable. To get a final prediction the score is thresholded and cases below the threshold are assigned a value of 0 and above the threshold to 1. Given a reference classification and output with a thresholded score we have four possible situations: (1) if the output and the reference are 1 we have a true positive (TP); (2) if the output and the reference are 0 we have a true negative (TN); (3) if the output is 1 and the reference is 0 we have a false positive (FP) and (4) if the output is 0 and the reference is 1 we have a false negative (FN). The performance is usually calculated considering these metrics (figure 1.7). We can define sensitivity (or TP rate) as the number of TPs divided by the sum of the number of TPs and FNs. The sensitivity shows the percentage of abnormal cases correctly identified by the classifier. The specificity is defined as the number of TNs divided by the sum of the number of TNs and FPs. It is a ratio that shows the percentage of normal cases correctly identified. Another metric related with specificity and that we use in ROC graphs is the FP rate defined as the number of FPs divided by the sum of TNs and FPs. Both metrics are directly related as $\text{Specificity} = 1 - \text{FP Rate}$.

Figure 1.7. Confusion matrix and common performance metrics. The hypothesized class will be the threshold score of the CAD system and the True class the reference labeled classes. (Fawcett 2006)

		True class		<div>$\text{FP Rate} = \frac{\text{FP}}{\text{N}}$$\text{TP Rate (recall)} = \frac{\text{TP}}{\text{P}}$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}}$$\text{F-measure} = \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}}$</div>
Predicted class				
	True Positives (TP)	False positives (FP)		
	False Negatives (FN)	True Negatives (TN)		
				<div>$\text{P} = \text{TP} + \text{FN}$$\text{N} = \text{TN} + \text{FP}$</div>

A ROC graph is a two dimensional representation with the TP rate in the Y axis and the FP rate plotted on the X axes. A ROC graph represents visually the trade-off in real classifiers between benefits (TPs) and costs (FPs). In ROC analysis a great variety of thresholds are calculated in order to get the pair sensitivity/specificity of each threshold. The representation of these pairs create the ROC curve. The closer the figure gets to the upper left corner (sensitivity of 1 and specificity of 1) the better the performance of the CAD system is.

Another global relevant ratios to determine the ability of a test to discriminate between healthy and unhealthy cases are the precision, the recall, the accuracy and the F-measure. Figure 1.7 contains the definition of this metrics. When dealing with highly skewed datasets, Precision-Recall (PR) curves gives an alternative to ROC curves (Davis and Goadrich 2006). In the PR curves, the x-axis represents the Recall and Precision on the y-axis. F-measure or F-score represents the balance of both previous metrics, precision and recall. The diagnostic accuracy or effectiveness of a test is defined as the proportion of correctly classified cases (TP+TN) among all cases (P+N).

Finally we should mention as a relevant performance ratio to evaluate machine learning classifiers, the Area under the ROC (AUC). This ratio represents the probability that a randomly chosen positive example is correctly rated with greater suspicion than a randomly chosen negative example (Bradley 1997). According to some studies it should be used AUC instead of accuracy to optimize and compare learning algorithms (J. Huang and Ling 2005).

Chapter 2

Objectives

As we have described in the previous chapter screening for disease represents a powerful tool in medical care. The design and implementation of a successful and cost-effective screening program must follow the guidelines and procedures mentioned. The adoption of new screening programs in diseases like NMSC and glaucoma are still in progress. We have reviewed studies that show its potential with promising outcomes in terms of accuracy in the diagnosis and cost-effectiveness, but they are still out of national screening programs. Telemedicine and image processing methods included in CAD systems are becoming a cornerstone in current screening programs. Tele-screening allows the access to early diagnosis and monitoring of multiple conditions to patients in remote locations reducing the cost of the final solution without compromising the quality of the service. CAD systems, on the other hand, represent a valuable tool in multiple relevant tasks (segmentation, detection, quantification of structures or findings, among others) used by the specialists to achieve an improved final diagnosis.

This PhD Thesis proposes a translational research that aims to outperform the technological tools included in screening programs in fields like NMSC and glaucoma. The research hypothesis that we intend to validate is that screening in NMSC and glaucoma can be improved with the creation and application of new advanced tools and image processing methods. In order to confirm this hypothesis we define as **the main objective of the present PhD Thesis the design, implementation and validation of new tools and image processing algorithms to be applied in screening programs of NMSC and glaucoma**. Although we focused on these conditions, all the tools, methodologies and algorithms implemented will eventually be applicable to other screening scenarios. In order to cover most of the technological components present in screening campaigns, we restrict the scope of the PhD works following the criteria included in the following table:

Table 2.1. Criteria scope definition for all the developments of the work.

Component	Criteria
Tele-screening	Multi-specialty covering at least NMSC and glaucoma
CAD – enhancement	OCT for NMSC
CAD – detection	Color fundus images and glaucoma

The two technological assets identified in screening services part of the research activities of the project are a telemedicine tool specialized in screening and specific algorithms to be included in CAD systems and specialized in certain tasks. A tele-screening application represents a basic component in order to facilitate the access to information needed in the remaining works, so it will be covered first. The implementation of tools to be part of CAD systems are considered in two different scenarios. We focus on skin OCT imaging for the enhancement tasks and in glaucoma with color fundus images for the detection tasks. Following this main objective and the restrictions mentioned we can divide the project in three specific sub-objectives:

1. Design, implementation and evaluation of a tele-screening tool.
2. Design, implementation and evaluation of an OCT enhancement algorithm for skin images.
3. Design, implementation and evaluation of a CAD system for glaucoma based on CNNs.

2.1 Tele-screening tool

As we have seen, the implementation of state-of-the-art screening programs take advantage of the facilities of store-and-forward telemedicine applications. In the context of this PhD thesis a tele-screening tool offers several advantages: (1) Provision of ICT services needed in the execution of a real screening program. (2) Help in the evaluation of the output of the different works performed during the project. (3) Structure of all the information collected during a campaign in order to create a labeled dataset (4) Integration with the CAD system that could provide a global solution to fulfil most of the technological needs of a successful screening service.

The following requirements were identified as targets of the tele-screening tool:

- Multi-specialty, multi-disease, multi-role and multi-center, in order to allow the use for different screening services and health conditions, but also for different types of centers (hospital, primary care centers, management centers...) and users (image technicians, physicians, clinical specialists, administrative personnel...).
- Coverage of all the main steps in a screening service: patient recruitment, patient examination and test collection and diagnostic.
- Secured web-based application, to enable the safe use of the system and to facilitate a global access from different devices and locations with minimum requirements.

The tool was deployed and extensively tested in a real screening program for eye diseases in collaboration with Hospital de la Esperanza of Barcelona (Spain). Although it is out of the scope of a generic solution, the deployment of the tool in a real scenario required the custom integration with the image acquisition devices and the ICT systems of the health centers where the tool was deployed.

2.2 OCT enhancement algorithm for skin images.

OCT has been extensively used in many clinical fields as we described in chapter 2.1.5. The new advances in hand held probes and the reduction in the cost of the new devices will allow the use of this technology in primary care settings and screening services for early diagnosis of diseases like eye pathologies or skin cancer. As we mentioned one of the first components in any CAD system is the pre-processing step. The enhancement of the image to be processed can aid the specialists for a better identification of structures and findings relevant for the diagnosis and also for automatic systems that must process the image to perform an equivalent task without human supervision. Speckle noise reduction in OCT is a very active research field that allows quantitative and qualitative improvements in the images.

To accomplish this objective we have identified the following targets:

- Revision of state-of-the-art speckle noise reduction algorithms in OCT.
- Design and implementation of a new speckle noise reduction method for OCT skin images.
- Acquisition of 3D OCT datasets to apply the different speckle noise reduction implementations.
- Quantitative evaluation of the output of the enhancement methods (including the new one) using typical performance ratios.
- Qualitative assessment of the output of the enhancement methods using ratios relevant for the early diagnosis of NMSC using a telemedicine framework.

These works were performed in collaboration with the Drexler Lab in the Center for Medical Physics and Biomedical Engineering of the Medical University of Vienna (Austria) that assisted in the acquisition of OCT datasets and in the clinical expert evaluation of the speckle noise reduction filters.

2.3 CAD system for glaucoma based on CNN.

The application of Convolutional Neural Networks, as in many other fields of computer vision and medical image processing, is a landmark in the field of CAD systems in recent years. The use of CNNs technology is especially attractive in the framework of a screening program due to the intensive use of images and their application in thousands of studies. We can consider a basic requirement to create a robust CAD system based on CNN, the acquisition of a relevant number of images labeled by specialists. The tele-screening tool and its application in a clinical screening program for eye disease provided enough data to cover this need.

To accomplish this objective we have identified the following targets:

- Revision of state-of-the-art CNN algorithms and available implementations.

- Collection and pre-processing of the labeled datasets used as gold-standard.
- Design, implementation and testing of a CNN algorithm to detect glaucoma using segmented color fundus images and examination data from the screened patients.
- Assessment of the new glaucoma CAD system using performance ratios and comparison with the performance of the evaluation of specialists.

These activities were performed in collaboration with the Diagnostic Image Analysis Group (DIAG) of the Radboud UMC of Nijmegen (The Netherlands) that has great experience in the field of Deep Learning and its application in CAD systems, and with Hospital de la Esperanza of Barcelona (Spain) to provide labeled datasets and clinical expertise.

2.4 Context

In this section we briefly describe the context of the research projects that supported the majority of the developments accomplished to better understand the objectives of the present work. Although the goals and motivations of these projects exceeded the scope of the present work, they could reveal important factors and requirements that influenced some decisions during the realization of the work.

The first one is the European Union FP7 project BiopsyPen¹. As we mentioned in chapter 1.1.5, the current gold standard practice for morphological evaluation of skin is the biopsy and the histopathological examination, an invasive method, expensive and complicated for the monitoring of dermatological lesions. In this sense, a non-invasive, high resolution and in-vivo imaging for early diagnosis of skin diseases like non-melanoma skin cancer (NMSC) could represent a great advance. Within this context, the main goal of the project was “to develop a broadly available, cost-effective reliable and compact optical biopsy tool which radically transform the capability of biophotonic technology for dermatology and point-of-care diagnosis”. The final device developed by the project included a pen-like Optical Coherence Tomography probe and an associate console to visualize and assess the images acquired by the system. A hand-held non-invasive system like the Biopsypen OCT device could provide great advantages in early diagnosis of NMSC for point-of-care and screening services.

The “Universidad Politécnica de Madrid” (UPM) was responsible of one the work packages of the project related with the design and development of signal and image processing technologies to enable the reconstruction of 2D and 3D OCT images with enough quality to allow a reliable diagnostic at point-of-care. One the most valuable contributions of our participation in the project were the image processing algorithms assessed for OCT image enhancement. The results and continuation of these works are aligned with the previously described objective 2.2 that will be developed in detail in chapter 4.

¹ <http://www.biopsypen.eu/>

The second project related with the present work was supported by the Spanish “Instituto de Salud Carlos III Fondo de Investigaciones Sanitarias” (Health Institute Carlos III, Health Research Fund¹) and was titled “Assessment of diagnostic precision, costs and utility of glaucoma screening based on image based detection, automated image analysis and telemedicine” (FIS PI15/00412). The motivation of the project in this case, was the high prevalence of glaucoma and the existence of effective methods for the treatment of the disease in early stages that could limit their advance and reduce the visual impairments that it produces. UPM was responsible of the deployment of a tele-screening tool to assist in the screening campaign and the design and implementation of image processing methods to aid in the automatic detection of glaucoma. As we can see, the objective 2.1 and 2.3 previously mentioned, are linked with the tasks performed in this project and will be described in chapters 3 and 5.

¹ <http://www.isciii.es/ISCIII/es/contenidos/fd-investigacion/fd-financiacion/convocatorias-ayudas-accion-estrategica-salud.shtml>

Chapter 3

Tele-screening tool

The telemedicine and e-health technology have been extensively used in screening programs all over the world in the last decades. The potential benefits of telemedicine: remote provision of care, improved access to information, services and care delivery or reduction of health-care costs, motivated the adoption of this technology in screening services of multiples fields like breast cancer, cervical cancer, diabetic retinopathy or skin diseases. In this thesis we will focus on the use of telemedicine for screening in pathologies with high prevalence of two medical specialties: ophthalmology and dermatology.

In the context of the project of the Health Institute Carlos III described previously in 2.4, the creation of a new tele-screening tool is proposed. The tele-screening tool has been successfully tested in two real scenarios. First, we accomplished a proof of concept through a quality assessment of OCT skin images. This implementation is described in chapter 4 in the qualitative evaluation section. Second, we deployed the tool in a screening campaign with more than 1000 patients for the detection of multiple eye pathologies. The application also served as basis for the provision of labeled data and clinical feedback for the implementation of the glaucoma detection method proposed in chapter 5. In this chapter we will describe the main requirements of the tool, the design principles, the technologies involved in the creation of the application and the main results obtained in the ophthalmology screening campaign. Although the campaign considered multiple pathologies we will pay special attention to glaucoma. We will revise also the basic concepts regarding the diagnosis of this pathology in order to understand the challenges and issues that we had to face in order to design the screening project and implement the tool. This description will also serve as a reference for the automatic classification of glaucoma, covered in chapter 5.

3.1 Introduction

Teleophthalmology is the branch of telemedicine that is responsible for providing diagnostic services and monitoring of ocular pathologies through the use of digital medical equipment and telecommunications technologies. Its benefits in costs and improvement in sensitivity and specificity have been highlighted in numerous projects for several year (Kanagasingam Yogesan et al. 2006). This field of the telemedicine has received an increasing interest in the last years due to the prevalence and significance of eye diseases and the lack of specialists in remote areas. Most of the studies that we can find in research are focused on one or two particular diseases at the same time and are centered in a broad range of issues like analysis of the clinical effectiveness, the cost effectiveness or the patient satisfaction.

In a comprehensive literature review (Bahaadinbeigy and Yogesan 2012) identified that the majority of the studies in teleophthalmology were focused on diabetic retinopathy, general ophthalmology, retinopathy of premature and glaucoma. The studies centered in general eye problems were targeted at confirming the benefits of teleophthalmology to assist practitioners. Other conclusions of the review were that the majority of the projects, more than 80%, were store-and-forward, where the data related with the screening is stored previously to their future use by an expert professional for their diagnostic evaluation. In this chapter we will focus only in the use of telemedicine to support screening of a particular disease or group of diseases and not in the use telemedicine to provide other type of services.

With the rising incidence and prevalence of diabetes and diabetic retinopathy (DR) the screening of this pathology is the most frequent among all the eye diseases (Ting, Cheung, and Wong 2016). Early detection and prompt treatment allow prevention of up to 98% of diabetes-related visual impairment (Ferris 1993). The tele-screening has reported to be cost-effective in DR screening in various countries (Kumar et al. 2005; Neubauer et al. 2003; Cuadros and Bresnick 2009). In other pathologies like Age-Macular-Degeneration or glaucoma the screening studies are in a preliminary state in comparison with DR. Some recent studies are showing the potential benefit of AMD screening applied in combination with DR screenings (Chan et al. 2015) and other publications present tele-glaucoma as an effective screening tool for glaucoma for remote under-services communities (Thomas et al. 2014) and in urban settings (Hark et al. 2017). But it is still an open research field where efficient telemedicine tools and CAD system can play a positive and significant role.

3.2 Glaucoma diagnosis

Before we start with the description of the design and implementation of the tele-screening tool we will describe briefly the basic concepts related with the diagnosis of glaucoma to better understand the challenges and issues of this complex disease.

As we mentioned in the introduction (section 1.1.4) glaucoma is a worldwide leading cause of irreversible vision loss that affects more than 70 million people (Quigley and Broman 2006). The glaucoma is a group of progressive neuropathies characterized by degeneration of retinal ganglion cells and resulting changes in the optic nerve head. The biological basis of glaucoma is not well understood and the factors that affect its progressions have not been fully characterized (Weinreb, Aung, and Medeiros 2014; Nickells et al. 2012). Glaucoma can be classified in two broad categories: open-angle glaucoma and angle-closure glaucoma. Although open-angle glaucoma is responsible of 74% of glaucoma cases (Quigley and Broman 2006), angle-closure glaucoma is more severe and produce the same number of irreversible visual loss (Day et al. 2012). The following tests are recommended to a proper diagnostic of glaucoma (“European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part 1 Supported by the EGS Foundation” 2017):

- Intraocular pressure (IOP) and tonometry. The IOP is the pressure created by the continual renewal of fluids inside the eye. The IOP could be estimated with a tonometer, a device which measures the pressure inside the eye in millimeters of mercury (mmHg). The mean IOP in adult population is estimated at 14-16 mmHg, with a standard deviation of 3.0 mmHg (Hoehn et al. 2013; Foster et al. 2011). Elevated IOP is considered a very consistent risk factor for the presence of glaucoma, although several studies found IOP lower than 22 mmHg in 25% to 50% of individual suffering glaucoma (Mitchell et al. 1996). So it is not an effective method as unique test for screening and should be complemented with other information. It has been demonstrated that the corneal thickness affects the measurement of the intraocular pressure (Bhan et al. 2002; Recep et al. 2001). We can include in the glaucoma evaluation a pachimetry test for measuring the corneal thickness and calculate the IOP with more precision.
- Gonioscopy. The purpose of this test is to analyze the angle configuration and assess the presence of angle closure. It is based on the recognition of angle landmarks and uses a simple handheld mirrored instrument placed on the patient’s eye followed by examination of the angle using a slit-lamp biomicroscope. This test is considered highly subjective, with poor reproducibility (Weinreb, Aung, and Medeiros 2014). Anterior segment optical coherence tomography and High-Definition optical coherence tomography represent other valuable alternative tests for imaging the anterior chamber angle and an aid in the diagnostic of angle-closure glaucoma (H.-T. Wong et al. 2009; Narayanaswamy et al. 2010).
- Optic nerve head and retinal nerve fibre layer. Glaucoma changes the appearance of the optic nerve head (ONH) and the retinal nerve fibre layer (RNFL) in a characteristic way.

In figure 3.1 we can distinguish the main retinal elements present in a color fundus image. For the clinical examination of the ONH and RNFL the features included in the following table should be assessed. All this features described in table 3.1 can be assessed using color fundus images although other imaging modality could be consider. RNFL layer defects can be assessed using Scanning Laser Polarimetry and OCT (Lim, Chattopadhyay, and Acharya 2012). The figure 3.1 present some typical structures present in color fundus images. In figure 3.2 we show example of the features of the ONH and the RNFL considered in a glaucoma examination.

Figure 3.1. Typical structures present in a color fundus image: optic disc, optic cup, neuroretinal rim and vessels (Haleem et al. 2013).

Table 3.1. Features of the optic nerve head (ONH) and the retinal nerve fibre layer (RNFL) to consider in a glaucoma examination. (Jonas, Budde, and Panda-Jonas 1999; Fingeret et al. 2005; “European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part 1 Supported by the EGS Foundation” 2017). Examples in figure 3.2.

Feature	Description
Neuroretinal rim	Glaucoma is characterized by progressive narrowing of the neuroretinal rim following different patterns.
Retinal nerve fibre layer	It is the earliest sign of glaucoma. RNFL appears as a bright bundle of striations unevenly distributed in normal eyes. The defects can consists on wedge-shaped dark areas with its tip touching the optic disc border or with diffuse RNFL defects that reduce the appearance of RNFL.
Optic disc haemorrhages	Indicate glaucoma progression and can be located on the retina, be just off the disc, in the margin of the retina and the optic disc or near to blood vessels. They are more common in normal tension glaucoma.
Parapapillary atrophy	Refers to the thinning and degradation of the chorioretinal tissue just outside the optic disc which is an indicator of progression of glaucoma. There are two types depending on the zone where they are present, alpha and beta types. The beta type is more important for glaucoma because is caused by atrophy of the retinal pigment epithelium It could be present also in normal eyes.
The ISNT rule	The rim width is the distance between the border of the optic disc and the position of blood vessels bending. In normal eyes the inferior area is usually wider than the superior followed by the nasal and the temporal the narrowest. The term ISNT helps to remember the anatomic rule. If the rim does not obey the ISNT rule glaucoma must be suspected.
Rim width and Cup/Disc Ratio	The width of the rim and the size of the cup vary with the overall size of the disc which is very different in the population. The vertical diameter of the cup and the disc is used to calculate the cup/disc ratio (CDR). A large CDR has been identified as a sign of glaucoma damage. However large CDR could be interpreted erroneously pathological in large discs and a small CDR in a small disc could be misclassified as normal. Another factor to analyze is that the cupping in normal eyes used to me symmetric, vertical CDR is less than 0.2 in 96% of the normal eyes.

Figure 3.2. Examples of features to be considered in the glaucoma examination of the ONH. Extracted from (“European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part 1 Supported by the EGS Foundation” 2017)¹. (a) Progression of glaucomatous damage of the optic disc: Early localized loss (A1), advancing to localized plus diffuse rim loss (A2). Early localized rim loss, polar notches (B1); more advanced polar notches (B2). Diffuse or concentric rim loss, early (C1); advanced (C2). Diffuse rim loss (D1), followed by localized rim loss (notch) (D2). (b) Optic disc haemorrhage. (c) ONH with parapapillary atrophy. The Alpha zone is located peripheral to beta zone and is characterized by irregular hypo- and hyperpigmentation. The Beta zone of atrophy is adjacent to the optic disc. (d) The ISNT rule. (e) Optic nerve heads with different disc areas but with the same rim area, and the same number of retinal nerve fibres; small size disc (disc areas less than 2 mm² and C/D=0.3), mid size disc (disc area between 2 and 3 mm², C/D=0.5) and large large disc (disc area greater than 3 mm² and C/D=0.8). C/D=CDR.

¹ Pages of the images in the document of the European Glaucoma Society. (a) 49 (b) 50 (c) 51 (d) 52 (e) 54

The quantitative imaging of the optic nerve and the RNFL are commonly used to assist in the diagnosis of glaucoma and the monitoring of the progression of the disease. The majority of the imaging devices have software that provides measurement parameters and a clinical diagnosis based on the statistical result of comparing the output measurements of the patient with a normative database of healthy eyes. The quality of the image, image artefacts or software errors are possible, so the clinicians should be careful interpreting this information and consider all the data and tests of the patient to perform a final diagnosis. Heidelberg Retina Tomography (HRT), Scanning laser polarimetry (GDx-ECC) and Optical Coherence Tomography (OCT) (Medeiros et al. 2004; Burgansky-Eliash et al. 2007; Weinreb et al. 1998; Jaffe and Caprioli 2004) have demonstrated to be valuable imaging technologies to assist in the detection and follow up of glaucoma.

- **Perimetry.** Visual fields tests are relevant for the diagnosis of glaucoma and even more important in the follow-up and management of the pathology. It is a subjective evaluation of central and peripheral vision and usually consists on light spots presented to the patient in different areas of the peripheral vision. Over the time the test has become more standardized and now it is usually named Standard Automated Perimetry (SAP) that refers to static computerized threshold perimetry of the central visual field. This tests are time consuming and required various repetitions in patient without experience (Kamdeu Fansi, Li, and Harasymowycz 2011) and have been reported in several studies that present a high rate of false positives (Pierre-Filho et al. 2006; Mowatt et al. 2008).

Although there are several comparative diagnostic studies, there is no evidence or agreement about the tests or combination of tests that should be performed in order to improve the benefits to the patient at a sustainable cost (Tuulonen 2011; Burr et al. 2007). The lack of a perfect reference standard and the great variability of technologies for establishing the diagnosis of glaucoma make early diagnosis and screening challenging and complex. In an extensive review (Ervin et al. 2012) conclude that there is no studies yet that reported evidence for direct or indirect links between glaucoma screening and visual fields loss, visual impairment, optic nerve damage, intraocular pressure, or patient-reported outcomes. As we can see, early diagnosis and glaucoma screening are an open and very active research field that requires further studies and new strategies to overcome current limitations. Some of the solutions proposed in the present work seek to advance in that direction.

3.3 Methods

The purpose of this section is the description of the process followed for the creation of a store-and-forward tele-screening tool that offered technological support for the implementation of multi-disease screening campaigns of multiple medical specialties. The following table summarize the main requirements identified in the design of the tele-screening tool.

Table 3.2. Global requirements of the tele-screening tool

ID	Description
REQ1	The application will cover all the steps in a screening campaign: recruitment of the patients, examination, remote evaluations of the study, integration with the medical record of the hospital and monitoring with posterior exams in the hospital or in future campaigns.
REQ2	Multi-center. The application will allow the use of different medical centers simultaneously related in a screening campaign.
REQ3	Multi-role. The application will allow the use of different roles associated with the process of the screening campaign: assistant, exam technician, evaluator, manager and administrator
REQ4	Multi-device. The application will be web-based allowing its use by different devices through an internet browser.
REQ5	Security. The platform will warrant the security in terms of authentication, confidentiality, integrity and in particular will apply the procedures established in the Spanish LOPD for the protection of the personal data of the users and patients of the platform.
REQ6	Cloud storage. The application will allow the secure storage of images, tests and reports involved in the campaign in a remote environment.

The architecture of the tele-screening tool followed the Model/Controller/View (MVC) design pattern, very common in interactive software systems and in web development (Leff and Rayfield 2001). In a MVC application the *View* displays the information to the end user, the *Controller* processes the user's interaction using the information store and organized in the *Model*.

The first step in the design of a MVC application is usually the creation of the *Model*. In our case this step supposed the design of a SQL relational database (Maier 1983) that includes all the information that could be processed by the tele-screening tool. We put special effort defining all the database tables and fields that must cover all the multiple options defined in the requirements of the tool, in particular the multi-specialty, multi-disease, multi-center and multi-role requirements. The following figure shows some tables defined in the database of the tool. We can divide the model in four modules responsible for different types of information.

- *Personal information module*, includes all the tables and fields related with the users of the tool, some personal data, the addresses, the roles, ...
- *Medical participants module*, includes all the tables for the patients, the health professionals the medical centers, the specialties, subspecialties, ...
- *Study-Evaluation module*, includes all the tables with the examinations fields, the tests, the diagnosis fields, ...
- *Administration module*, with all the tables and fields regarding the agenda, the logs, the task assignment, ...

Figure 3.3. Diagram of the modules of the database model of the tele-screening tool.

As we mentioned in the requirements, the tool is multi-role (REQ3 in table 3.2). One of the tables in the model stores all the roles considered which are described in the next table:

Table 3.3. Roles of the tele-screening tool.

ID	Description
admin	Access to all the functionalities of the application and the management of the users, the centers and the import/export facilities.
assistant	Management of the agenda and the patients of the campaign.
requester	Responsible for the creation of the study that includes clinical data, tests and reports performed during the examination of the patient.
evaluator	Responsible for the assessment and diagnosis of the studies based on the identification of pathological findings from the analysis of all the information upload by the requester in the examination.
manager	Access to the information of the schedule, the results of the examinations and the evaluation and the statistics of the campaign.

Next we defined the *Controller* where we identified the different functionalities that must be processed by the tool. Although the tool was not designed as a service oriented architecture (Erl 2005) we tried to include some of the ideas and concepts of this powerful software paradigm in terms of abstraction of logic and reusability. Additionally we design the logic of the application with the vision of the process followed in a typical screening campaign represented in the figure 3.4. Briefly, we can divide the screening process in different activities included in three main groups: (1) the agenda sub-process that involves the contact with the patient and the schedule of the examination in the system, (2) the examination sub-process that covers all the activities related with the realization of the study to the patient and the inclusion of all the data in the system. Finally (3) the evaluation sub-process that accounts for all the tasks related with the diagnostic evaluation and the integration of the information with the Electronic Medical Record (EMR) system of the hospital.

Figure 3.4. Screening process supported by the tele-screening tool.

The next figure present the main functionalities and the modules of the tele-screening tool. We describe briefly the main modules present in the figure that cover the core functionalities accomplished by the tool:

- The *agenda* module includes all the functionalities related with the management of the examination scheduled for every date.
- In the *exam* module all the issues related with the incorporation of data of the medical history, results of tests, reports or images needed for the posterior evaluation were implemented.
- The *allocator* module deals with the assignation of studies to the different evaluators included in the screening campaign. It also allows re-assignation in case of double evaluation, especially interesting in research projects.
- The *evaluations* module includes all the functionality regarding the assessment of the studies previously included. The evaluators must indicate all the findings related with the different diseases of the screening campaign. The tool provides a reference diagnosis according to the selection of findings of the evaluator that could be changed at the end of the process.
- The *report diagnosis* module is responsible of generating the report with all the information of the study, including all the examinations, images, the clinical diagnosis of the evaluator and the next steps if any. This report could be incorporated to the EMR system of the hospital if needed.

Figure 3.5. Functional modules of the tele-screening tool. The integration with the Electronic Medical Record (EMR) of the Hospital although included in the global design of the platform required ad-hoc developments in each deployment due to the variability of the IT systems in hospitals.

Finally, for the *View* we identified and prototyped all the different windows that should be present in the application to comply with all the requirements described before. In the next figures we present all the basic views created for the tele-screening tools and two examples of the final views implemented for the ophthalmology campaign. It is important to highlight that the views were defined for a specific application of the tool in an ophthalmology screening campaign but the approach and design of the tool allows the creation of similar views for different implementations in campaigns of other medical specialties and focused in other diseases.

Figure 3.6. Example of views included in the tele-screening tool for the roles admin, study requester, evaluator and manager. As we can see some of the views are repeated but could be not exactly the same because the visible information for each role could be different. Also it is important to mention that in most of the cases each view in the figure could include several windows for the final user. For example the study creation involves 3 windows.

The implementation of the tele-screening tool used typical web-based open software technologies. The most relevant are presented in the following table.

Table 3.4. Software packages used in the implementation of the tele-screening tool.

Name	Version	Reference
Laravel Framework	4.2	PHP open source Framework (https://laravel.com/)
Bootstrap	3.3.2	HTML, JS, CSS open source toolkit (https://getbootstrap.com/)
MySQL	5.6	Open-source relational database management system (https://www.mysql.com)
Imagick	3.1.2	PHP Imaging Processing Library (http://php.net/manual/en/book.imagick.php)
Chart JS	1.0.2	Javascript charting library (https://www.chartjs.org/)

One benefit of the web-based strategy with cloud storage of the images is that it simplifies the requirements for the installation of the tool in a real environment. The inclusion of the agenda of the studies to be performed and the integration of the final report with all the

diagnostics with the Electronic Medical Record of the hospital are the unique custom developments that need to be implemented. In the next section we will detail the solution included in the ophthalmology screening study for this purpose.

3.4 Results and discussion

For the evaluation of the tool we completed its implementation in a real ophthalmology screening campaign performed in the Hospital Esperanza and the Primary Care Centers of Vila Olímpica and Barceloneta in Barcelona (Spain). As we have mentioned before, the tele-screening tool allows the study of multiple diseases at same the same time. In this case we included eye diseases described in the next table. In the rest of the descriptions we will focus mainly on glaucoma which is our target disease in this part of the work.

Table 3.5. Eye pathologies included in the tool for the screening campaign performed in the Hospital Esperanza in Barcelona (Spain)

Pathology	Types included
Diabetic Retinopathy (DR)	DR non proliferative (mild/moderate/severe) (with macular edema mild/moderate/severe) or DR proliferative (with macular edema mild/moderate/severe)
AMD	(mild/intermediate/severe)
	Age Physiological changes
Glaucoma	Suspect of glaucoma
	High intraocular pressure
Others	Vitreoretinal interface changes (vitreomacular adhesion / vitreomacular traction/macular hole/ epiretinal membrane)
	Vascular obstructive disease
	Retinal Photocoagulation

Next we describe the process followed during the campaign and the support provided by the tele-screening tool. The next figure summarizes all the steps and the number and role of the users of the tool involved:

Table 3.6. Process followed during the screening campaign. In each step the number of the users and the role are indicated (#users/role).

- First, the recruitment and scheduling of patients was performed. The staff of the Primary Care Center called and scheduled the visits of the people that agreed voluntarily to be

included in the study. From an initial random selection of 4000 patients of age older than 55 years old 1644 accepted to be included in the study and were scheduled in the agenda. To aid in this process the tool included a module specially implemented for this campaign that allowed the automatic import of all the patients scheduled. Besides the personal information of the patient, some data from the clinical history were also provided (diabetes, hypertension, hyperlipidemia, cerebral ischemia, smoking habits as well as levels of glomerular filtration and glycosylated hemoglobin). In the next table we present the age and gender distribution of the patients that finally accepted to participate in the study.

Table 3.7. Age and gender distribution of patients included in the study. In brackets the percentage respect to the TOTAL number of patients (1006) is included. The median of the age of the participants was 64 years old and the standard deviation 8 years.

Age	Number (%)	Gender	Number (%)
≤ 65	423 (42.05)	Female	523 (51.99)
≤ 75	370 (36.78)	Male	483 (48.01)
≤ 85	198 (19.68)		
> 85	15 (1.49)		

- The next step was the examination of the patients. Seven health professionals (nurses and optometrists) were responsible of performing medical tests and uploading the results in the telemedicine tool. The tests performed and the data collected and uploaded to the tool are listed in the following table. Once the study is completed, the tool automatically sent the study to a pool of eight ophthalmologists

Table 3.8. Data and tests included in the examination of the patients and uploaded to the tele-screening tool. AMD, age-related macular degeneration.

Data/Test	Description
General Clinical Data	Diabetes, Diabetes treatment, HbA1c, Glomerular filtration rate, Hypertension, Hypertension Control, Cerebral Ischemia, Ischemic heart diseases, Smoker habits
Ophthalmological Clinical Data	Diabetic Retinopathy/Glaucoma/AMD/ Family History Glaucoma/Ocular Hypertension/Glaucoma Layer Therapy/Diabetic Retinopathy/AMD/Retina Laser therapy/Retina Therapy/ Cataract Surgery/Refractive Surgery Personal History
Visual Acuity	With pin hole or with correction.
Tonometry	For measuring the IntraOcular Pressure (IOP) using a portable tonometer Icare.
Fundus Color Images	One for each eye with a field of view of 45 degrees, centered on the macula and including the optic disc using a Non-Mydriatic Retinal Camera Topcon
OCT	Three reports, one of the papilla of both eyes and others from the macula of each eye using the iVue 100 by Optovue, Inc Fremont, CA, USA.

In the tool the uploading of all this information was divided in three sections, the first one for the general clinical data, the second for the ophthalmological clinical data and the last one for the tests. In the next figure we can see an example of the last view that the user had to complete to finish the study.

Figure 3.7. Example of view of the last step in the study creation. In the top part of the image the user can enter the results of the tests for the visual acuity and the tonometry. Next, the user can upload the image tests acquired. In this case we can see two color fundus images. The user in any moment can save the study, go back to the previous step (the ophthalmological clinical data) or finish the study and allow the tool to send it to one evaluator.

- Next, the specialists evaluated the study using all the information previously uploaded. The assessment consisted in labeling all the findings identified through the data of the study. Figure 3.8 shows an example for the glaucoma examination. Once the evaluator completed the labeling, the tool automatically classified the pathology according to a protocol adapted from the recommendations of the European Glaucoma Society (“European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part 1 Supported by the EGS Foundation” 2017) and presented in figure 3.9. The ophthalmologist can modify the final classification before concluding the diagnosis. This stage ends with the creation of a PDF report (figure 3.10) that was sent to the Health

Center for its later delivery to the patient and with the scheduling in the Hospital's agenda if it is identified that the patient has a positive diagnosis.

Figure 3.8. Example of the evaluation view in the assessment of glaucoma. In the image we can see a zoom of the optic disc of one of the color fundus images and the fields that the evaluator has to label: rim thinning, defect nerve fiber layer, rule ISNT, peripapillary atrophy beta type, papillary haemorrhage and the vertical index papilla excavation (similar to the CDR that we described). We can contrast that these fields cover all the features identified in table 3.1 which are recommended to assess the ONH and the RNFL.

Dyseo - Main Environment Juanjo Gómez (Evaluador) Exit

1- Fundus photograph 2- Fundus photograph 3- OCT Optic Disc

Download Zoom

Type: Fundus photograph Date: 19-07-2018 Eye: Right Position: NM1 - Macula and Optic Disc

Description:

Evaluation

Repeat in:

Retinal interface **Glaucoma**

No signs suspect glaucoma	☐ RE	☐ LE	No evaluation	☐ RE	☐ LE
Rim thinning	☐ RE	☐ LE	Rule ISNT breach	☐ RE	☐ LE
Defect nerve fiber layer	☐ RE	☐ LE	Papillary hemorrhage	☐ RE	☐ LE
Peripapillary atrophy beta type	☐ RE	☐ LE			

Vertical index papilla excavation

RE LE

Observations:

Figure 3.9. Protocol followed in the tele-screening tool to identify the glaucoma suspects based on the features labeled by the evaluator in the view of the figure 3.8. CDR is the cup/disc ratio or vertical index papilla excavation as indicated in the view. The $\Delta\text{CDR}_{\text{LRE}}$ is the difference of the values selected by the evaluator in the Right and Left Eye (LRE).

Figure 3.10. Example of report created by the tele-screening tool. The report contains information of the study, the tests performed, the assessment of all the pathologies and of each test and the next steps. In the figure we can see the page with diagnostic for all the pathologies considered in the study and the page with the evaluation of each test and the recommendations.

Detection Program Ophthalmological Diseases
Hospital de l'Esperança

Evaluation Report

Diabetic Retinopathy

Right Eye: No signs of diabetic retinopathy
Left Eye: No signs of diabetic retinopathy
Observations:

AMD

Right Eye: No signs of AMD
Left Eye: No signs of AMD
Observations:

Alterations vitreo retinal interface

Right Eye: Non evaluated
Left Eye: Non evaluated
Observations:

Glaucoma

Right Eye: Suspect glaucoma
Left Eye: No signs of suspect glaucoma in the images
Observations:

Other Pathologies

Right Eye:
Left Eye:

2 / 3

Detection Program Ophthalmological Diseases
Hospital de l'Esperança

Test/Reports

1 - Fundus photograph, Right, 19/07/2018.
Quality: Good. State: Normal.
Recommendations:

2 - Fundus photograph, Left, 19/07/2018.
Quality: Good. State: Normal.
Recommendations:

3 - OCT Optic Disc, Both, 19/07/2018.
Quality: Good. State: Suspect. Diagnostic: Suspect glaucoma.
Recommendations:

Recommendations

New Revision: You will be notified by your Health Center
Revision Type and Center: Ophthalmological Campaign Consultation. Hospital de l'Esperança

3 / 3

- Finally, the last step in the screening process has been the hospital complete examination of all the patients with a positive diagnosis and part of the negatives following the criteria of two negatives for each positive (with the same age and sex). In the case of glaucoma the hospital evaluation includes a detailed exploration of the anterior and posterior segment, a gonioscopy test, a pachymetry test and a perimetry test. 22 users, 11 requester and 11 evaluators are introducing the data of this tests in the tele-screening tool. During the writing of this document this last step has not been finished yet.

The average time of performing the tests and loading in the telemedicine tool has been determined between 12 and 15 minutes. The evaluation of the data and images of each patient and the preparation of the report requires between 1 and 3 minutes for the specialist and no significant difficulties have been detected in the evaluation process.

In table 3.9 we present the total number of cases with one of the main diseases considered in the campaign: DR, AMD and glaucoma. The following numbers show a comparatively low number of positives cases of DR compared with AMD and glaucoma. One possible explanation to this difference is that the diabetic patients perform regular controls and were not very relevant in the study. The figures in the case of AMD and Glaucoma, two pathologies with high prevalence, show an important number of suspects that must be confirmed in the final evaluation in the hospital. The final number of pathological cases are expected to decrease because the prevalence of these diseases are lower than these figures. The prevalence of early AMD, for example in Europe, depending on the age ranged from 3.5% in those aged 55-59, to 17.6% in those aged >84 (Colijn et al. 2017). In the case of glaucoma the prevalence for population aged 40-80 is estimated in 3-4% (Tham et al. 2014).

Table 3.9 Age-related Macular Degeneration (AMD), glaucoma and Diabetic retinopathy total diagnosis. In brackets we present the percentage respect to the total number of studies: 1006.

Diagnostic	Number of patients (%)
Age-related macular degeneration	154 (15.31)
Glaucoma	179 (17.79)
Diabetic Retinopathy	31 (3.08)

If we focus on glaucoma, in the next tables we present the features that were labeled in the glaucoma evaluation by the ophthalmologist. The majority of the cases were related with the identification of the features that we described previously (table 3.1). Tables 3.10 and 3.11 show in more detail which of these findings were more relevant. The violation of the ISNT rule with 80 cases was the most labeled feature followed by the rim thinning with 63. In 36 cases both features were identified together. It is also remarkable that only for 6 patients the only motivation to classify the study with a glaucoma diagnosis was due to a high IOP, that besides has to be confirmed in the hospital review. In the rest 173 cases other findings

were identified. In our study the relevance of this test could be considered limited. A similar recent study performed in Philadelphia (Hark et al. 2017) found a higher percentage of ocular hypertension, 6.1% and of suspicious optic nerve appearance 28.5%. The comparison has some limitations because in this study 61% of the participants were african-american which considered a glaucoma risk factor. Another important difference is that they did not used OCT in the first examination.

Table 3.10 Glaucoma diagnosis and detail of the evaluation labeling. The pathological findings in the ONH considered are: rim thinning, defect nerve fiber layer, rule ISNT violation, peripapillary atrophy beta type or papillary haemorrhage. The CDR is the cup/disc ratio and $\Delta\text{CDR}_{\text{RLE}}$ is the difference of the CDR in both eyes. In brackets we present the percentage respect to the total number of studies: 1006.

Diagnostic	Number of patients (%)
Total Glaucoma	179 (17.79)
IOP $\geq 22\text{mmHg}$	27 (2.68)
Pathological finding in ONH assessment	147 (14.61)
CDR ≥ 0.8	25 (2.49)
$\Delta\text{CDR}_{\text{RLE}} \geq 0.3$	19 (1.89)

Table 3.11 Detail of the evaluation findings in the ONH. Total number of studies 1006.

Diagnostic	Number of patients
Total Pathological findings in ONH	147
Rim thinning	63
Defect nerve fiber layer	34
Rule ISNT violation	80
Papillary haemorrhage	6
Peripapillary atrophy beta type	52

Other relevant output of the results is the analysis of the tests used in the campaign. In table 3.12 we present the image quality assessment by the evaluators. We can highlight that color fundus images had a significant inferior image quality than OCT. The total number of color fundus images that were not evaluable or had a bad quality represent 8.58% of the total, that compared with other similar studies represent an intermediate percentage between the 5% reported in (Ahmed et al. 2013) and the 24 % of (Kiage et al. 2013).

Table 3.12 Detail of the evaluation of the quality of the images and the percentage related with the total number of images. Total number of studies 1006.

Diagnostic	Total	Good (%)	Enough (%)	Bad (%)	Not evaluable (%)
Color Fundus Images	2049	1432 (68.69)	441 (21.52)	144 (7.03)	32 (1.56)
OCT macula	1991	1863 (93.57)	96 (4.82)	19 (0.95)	13 (0.65)
OCT Optic disc	994	920 (92.56)	47 (4.73)	11 (1.11)	16 (1.61)

Finally we show in table 3.13 the specific classification performed by the evaluators in the analysis of each image test provided in the study. The assessment of the ONH in the color fundus images and the OCT optic disc report represent the most relevant tests to classify glaucoma. The last column shows the number of cases where the glaucoma diagnosis was based on the findings in only one test. We can see that color fundus and OCT of the optic disc detected an important number of cases individually. Until the confirmation of the final numbers of the hospital review of the patients, we can highlight that 44 patients out of the 179 (24.58%) were diagnosed thanks to the OCT, which represent a remarkable increment of suspects of glaucoma.

Table 3.13 Glaucoma diagnosis by image type. The unique test represents the cases identified only with this image test.

Diagnostic	Glaucoma	Glaucoma Suspect	Unique test
Color Fundus Images	38	120	50
OCT macula	3	12	1
OCT Optic disc	41	64	43

Tele-screening for eye diseases like glaucoma is an open research field where studies like the one presented in this chapter, could be a valuable asset to a better assessment of the benefits of these approaches in terms of patient outcomes at a sustainable cost. One of the key issues of this type of projects is the selection of the tests for the examination. One of the core elements in our study has been the inclusion of OCT and not the visual fields tests in the examination. The inclusion of an OCT in the examination as a complementary test and the long time required to obtain the visual fields tests motivated this selection. The use of OCT in glaucoma tele-screening is limited (Kassam et al. 2013; Arora, Rudnisky, and Damji 2014) and has been tested in conventional screening without yet conclusive benefits (Bussell, Wollstein, and Schuman 2014). The possible limitations in terms of sensitivity in screening using OCT could be overcome with the inclusion of the color fundus image assessment. As we have seen in the previous data (table 3.15) both tests were determinant in detecting suspicious cases independently. As we have mentioned the qualitative assessment of the ONH in color fundus images is a challenging task. This is one of the motivations of the numerous automatic detection algorithms that have been proposed and we will discuss in detail in chapter 5.

3.5 Conclusions

In this chapter we have presented the design, full implementation and evaluation of a store-and-forward web-based tele-screening tool. The design was based on the MVC design pattern and the implementation used web-based open software packages. The main output of the proposed tool is the technological transfer and application in an extensive study accomplished through an ophthalmology screening campaign performed in Hospital Esperanza in Barcelona that included 1006 subjects. The tool supported successfully all the requirements of the study and showed the relevance of this type of campaigns in highly prevalent pathologies like AMD and glaucoma where the number of cases with a pathological evaluation was high compared with other pathologies like diabetic retinopathy where the monitoring of the diabetic patients showed inferior positive cases. As we will see in chapter 5 the development of CAD systems that help in the diagnosis could facilitate the deployment of new screening programs for pathologies like AMD or glaucoma.

Chapter 4

Enhancement for OCT skin imaging

We present a new speckle noise reduction method for OCT volumes. The algorithm exploits the advantages of the adaptive noise wavelet thresholding proposing an improved compounding method of several frames at the same position. The proposed method is compared quantitatively with different state-of-art alternatives through its application to four different clinical OCT volumes of 256 acquisitions using a SS OCT specially designed for skin imaging. We additionally present a novel scheme for qualitative assessment of the results of the different techniques compared, that was performed by one specialist in the evaluation of skin OCT images. Both quantitatively and qualitatively the proposed method demonstrated the best performance, showing its possible application in the reduction of noise for a better diagnosis in OCT skin imaging or in preprocessing steps for posterior segmentation or quantification algorithms included in CAD systems for screening campaigns.

As we mention in section 2.4 this development was part of the activities preformed in the project BIOPSYPEN were we collaborated with the Center of Medical Physics and Biomedical Engineering of the Medical University of Vienna with the supervision of Professor Wolfgang Drexler director of the Center. Zhe Chen and Elisabet Rank, researchers of the Center, helped us in the acquisition of datasets and the dermatologist Christoph Sinz performed the qualitative assessment. We sincerely thank all of them for their help and assistance during the realization of this study.

4.1 Introduction

Optical coherence tomography (OCT) (D. Huang et al. 1991) is an optical imaging technique that allows cross-sectional views of highly scattering biological in vivo tissue in real time with micrometer resolution and depth up to several millimeters. OCT is broadly used in multiple biomedical and clinical fields, such as in ophthalmology as a routine noninvasive tool for the diagnosis and monitoring of disease progression [2] or in cardiology as a catheter-based imaging system for the guiding in coronary interventions (Bezerra et al. 2009). In dermatology OCT is a useful non-invasive tool that has been used to study a variety of dermatological disorders like systemic lupus erythematosus (Gambichler, Hyun, et al. 2007), bullous diseases (Mogensen et al. 2008), psoriasis and contact dermatitis (Morsy et al. 2010; Welzel, Bruhns, and Wolff 2003), wound healing (Greaves et al. 2014) or vascular skin lesions (Salvini et al. 2008). One of the most prominent application of OCT in dermatology is the assessment of non-melanocytic epithelial tumors: actinic keratosis (AK) (Mogensen et al. 2009; Jennifer K. Barton et al. 2003), basal cell carcinoma (BCC) (Mogensen et al. 2009; Gambichler, Orlikov, et al. 2007; Hussain, Themstrup, and Jemec 2015; Olmedo et al. 2006) and squamous cell carcinoma (SCC) (Hatta et al. 2010).

Due to the use of spatially coherent illumination OCT images are affected by speckle phenomena that has a dual role as source of noise and carrier of information (Schmitt, Xiang, and Yung 1999; Bashkansky and Reintjes 2000). The speckle carrying signal information is the result of the back scattering of the incident photons while the speckle noise is caused by the random interference between multiple reflected photons coming from multiple directions. Speckle noise creates a grainy appearance in the OCT images that degrades the signal-to-noise (SNR) and limits its interpretation. Their properties are affected by the scale representation, the optical setting and the scattering properties of the biological tissue (P. Lee, Gao, and Zhang 2011). As a source of information the speckle pattern statistics have been used to detect blood flow (Mariampillai et al. 2008; Srinivasan et al. 2010), differentiate living and non-living tissue (Gossage et al. 2006) or to perform optical elastography (Kirkpatrick, Wang, and Duncan 2006).

Speckle denoising is an active research field broadly developed during the last years. The different methods can be divided in those that modify the image configuration, where the optical setting or the scanning protocol can be adjusted, and those based in post-processing the images using digital algorithms. In the first group we can find different forms of compounding techniques using the frequency (Pircher et al. 2003a), the angle of the incident light source (Iftimia, Bouma, and Tearney 2003; Desjardins et al. 2006) or multiple A-lines collected in a controlled way (Szkulmowski et al. 2012). Recent studies propose methods that combine image processing and modifications in the optical configuration to increase SNR even further (Duan et al. 2014; W. Cheng et al. 2017). To compensate the effect of possible shifts in case of in vivo imaging two strategies have been proposed, either using a weighted average of individual A-scans (Sander et al. 2005) or a combination of high-speed

CMOS line scan camera and a cross correlation of different B-scans to a fixed reference one (Götzinger et al. 2011). One of the latest works suggests that the SNR improvement of pure angular compounding techniques are limited by optical aberrations (Winetraub et al. 2017).

The post-processing digital speckle reduction techniques have as a main advantage that they can be applied to almost all 2D and 3D images acquired by an OCT device without changing the acquisition setup. The drawback is that they usually need higher computation requirements and can affect the resolution of the image. We can include in this group multiple methods like local averaging over neighboring A-scans of each tomogram (Pircher et al. 2003b), averaging multiple B-scans (Sander et al. 2005), applying rotation kernel transformations to each tomogram (Rogowska and Brezinski 2000), image regularization (Marks, Ralston, and Boppart 2005) or complex diffusion filtering (Fernández, Salinas, and Puliafito 2005). Algorithms based on filtering in the wavelet domain have shown excellent performance for speckle noise removal. The basic approach of these methods is to filter the wavelet coefficients on multiple resolutions in order to minimize the noise. The calculation of this threshold and its application can be performed using spatially adaptive soft-thresholding with an estimation of the noise in one sub-band (Adler, Ko, and Fujimoto 2004; Chitchian, Fiddy, and Fried 2009). However, a recent study shows that speckle noise has different magnitudes at different wavelet sub-bands and that an estimation of the noise variance in individual scales can improve the characterization of the threshold and outperform the speckle noise removal (Zaki et al. 2017). Another reported successful strategy is the compounding of several frames in combination with digital filtering (Ozcan et al. 2007) and more specifically with frames previously filtered using wavelet denoising (Mayer et al. 2012).

We present a new method for speckle reduction of multiple OCT frames previously processed with a multi-scale noise adaptive wavelet filter before compounding. Similarly to (Zaki et al. 2017) we estimate the noise variance wavelet representation in a homogeneous scattering sample and use it to filter the noise in all the frames. To take advantage of complementary information brought from several acquisitions as a second step we used the compounding approach of (Mayer et al. 2012) applying a weight to select the detail coefficients of each sub-band before compounding all the filtered frames. We take advantage of the fact that the wavelet representation of the speckle statistics calculated from a homogenous sample can be used in other OCT images. These reference statistics are a valuable asset to differentiate information from speckle noise. Additionally the use of several frames and the wavelet weighted compounding allow us to get an improvement in the speckle removal and enhance the structure details showing better performance with respect to other methods.

The rest of the chapter is structured in several sections. In section 4.2 we describe in detail the proposed method, the OCT imaging system used, the data set, and the performance metrics used in the assessment. In section 4.3 we report the result of the quantitative and

qualitative evaluation of the algorithm. Finally, section 4.4 contains the summary of the study with the main conclusions.

4.2 Method

4.2.1 Wavelet Compounding Adaptive Noise Filter (WCAN)

Figure 4.1 shows a complete overview of the proposed method named WCAN. The input of the algorithm is a set of several noisy OCT frames ($OF_1, OF_2 \dots OF_N$) and a set of reference OCT frames of a homogeneous scattering sample ($RF_1, RF_2 \dots RF_M$). All the input frames are in logarithmic scale to allow the assumption of an additive noise model and normalized to 1 as maximum value. The output of the method is a single OCT denoised image. The minimum number of frames of each set are $N=2$ and $M=1$. The terms “frame” and “image” are used indistinctly in all the text, like the terms “sub-band” and “sub-scale” when we describe the wavelet decomposition. In the next paragraphs we explain in detail each step of the method.

Figure 4.1. Flow diagram of the wavelet compounding adaptive noise algorithm. The m initial OCT frames $IF_1, IF_2, \dots, IF_N$ could be previously digital filtered, all the images and the n reference OCT frames $RF_1, RF_2, \dots, RF_M$ are wavelet decomposed in L sub-bands. The variance of all the sub-bands of the wavelet decomposition of the reference stack is computed. Using the sub-band variances the weights of the detail coefficients for each sub-band is calculated and applied to each original OCT frame. Then, all denoised frames are compounded in the wavelet domain before the inverse wavelet transform is calculated to get the final denoised OCT image.

Pre-processing with digital filtering: A first step of pre-preprocessing with digital filtering could be included at the beginning to facilitate the posterior compounding. Empirically we identified the Sigma filter, also known as Lee Filter (J.-S. Lee 1983), as the best option for this step. The Sigma filter is based on the two-sigma probability of Gaussian distribution and incorporates the speckle multiplicative noise model. Besides, its simplicity it provides a good balance between filtering accuracy and computational complexity. The Lee Sigma Filter is based on the idea that 95.5% are distributed within the two-sigma range from its mean (J. S. Lee et al. 2009). We implemented a version of this filter considering a window size of 3 and the estimation of the sigma value based on the noise standard deviation obtained as the median absolute deviation of the diagonal detail coefficients sub-band of the single discrete 2-D wavelet decomposition of the input images using wavelet Daubechies 1 as mentioned in (Pizurica and Philips 2006)

Wavelet Transform: All the input images are decomposed by a wavelet transformation with a maximum decomposition level of L. The result of the transformation is a set of approximation coefficients A_i^l and detail coefficients $D_{i,O}^l$ where i is the frame number, l is the decomposition level and O is the orientation or direction of the detail coefficient (horizontal, vertical and diagonal). The wavelet transformation was computed using the 2-D discrete stationary wavelet transform with Haar wavelets.

Variance Computation: For the calculation of the noise variance (σ^2) in all the frames we consider all the sub-bands of the input and the reference frames. The variance of the reference frames is calculated once. For the computation of the variance of the detail coefficients of each sub-band all frames (M for the reference stack and N for the original stack) are taken into account.

We compute $\sigma_{R,D,M}^l$ as the set of variance values, where R refers to the reference OCT frames, D refers to the detail coefficients that can be in the horizontal, vertical and diagonal direction, l refers to the sub-band level (with l between 1 and L), and M is the total number of frames for the references stack.

Weight Computation: The weights provide an estimation of the noise contribution of each sub-band in every initial frame in relation with the reference noise and the rest of the frames. The goal is to limit the detail coefficients in each sub-band with higher variance compared with the reference values and the rest of the frames, assuming that the noise will be the main cause of this higher contribution. For the calculation of the weights we used the following expressions:

$$G_{i,D}^l(x) = \frac{1}{N} \sum_{j=1}^N \left| D_j(x) \cdot \frac{\sigma_{O,D,j}^l}{\sigma_{R,D,M}^l} \right| \quad (1)$$

$$T(\sigma_{i,D}^l) = \frac{\sigma_{R,D,M}^{l^2}}{\sqrt{|\sigma_{O,D,i}^{l^2} - \sigma_{R,D,M}^{l^2}|}} \quad (2)$$

$$W_{i,D}^l(x) = \begin{cases} 1; & |D_i^l(x)| > k \cdot T(\sigma_{i,D}^l) \\ 1 - G_{i,D}^l(x); & \text{otherwise} \end{cases} \quad (3)$$

where $G_{i,D}^l$ in (1) is the weighted average of all the detail coefficients $D_j(x)$ using the variance of the original frames in the sub-band l , $\sigma_{O,D,j}^l$, with respect to the variance of the reference frame $\sigma_{R,D,M}^l$. $T(\sigma_{i,D}^l)$ in (2) is an adaptive variance threshold defined in a similar way to standard wavelet thresholding (S. G. Chang, Yu, and Vetterli 2000) but considering the variance of the reference frames in each sub-scale $\sigma_{R,D,M}^l$. The expression (3) will produce the final weights $W_{i,D}^l$ for each sub-band l , frame i and detail coefficient D , comparing the value of the coefficient with the previous threshold value. The parameter k in (3) will be used to balance the final amount of noise reduction.

Wavelet coefficient weighting, compounding and inverse wavelet transform: The new detail coefficients of each sub-band $\tilde{D}_i^l$ are computed using the previous weights through the following expression:

$$\tilde{D}_i^l(x) = D_i^l(x) \cdot W_{i,D}^l(x) \quad (5)$$

The approximation and the detail coefficients of the denoised image are calculated averaging the coefficients of the N frames considered during the process:

$$D^l(x) = \frac{1}{N} \sum_{i=1}^N \tilde{D}_i^l(x) \quad (6)$$

$$A(x) = \frac{1}{N} \sum_{i=1}^N A_i(x) \quad (7)$$

The final denoised image is computed applying the inverse wavelet transform over the averaged coefficients.

The implementation of the algorithm was done in Matlab 2016b (Mathworks, Inc., Natick, USA) on a personal computer (Intel 3.3 GHz CPU, 32 GB memory). The final parameters used in the implementation are presented in the next table.

Table 4.1. Parameters used in the implementation of the method.

Parameter	Value	Description
M	2048	Number of frames for noise variance estimation (calculated once)
N	4	Number of initial frames at the same position
L	4	Maximum wavelet decomposition level
k	1.5	Adjustment of the thresholding and final noise reduction

4.2.2 OCT Imaging System

We used one OCT imaging system to evaluate the method. The tests with a custom state-of-art device provided a valuable feedback of the technical performance of the method. In addition we performed a qualitative assessment made by expert dermatologists, the potential and benefits of the use of this type of filters in usual clinic situations.

Swept Source OCT: The system used to evaluate the proposed method is a custom Swept Source OCT with a kinetic source from the Center for Medical Physics and Biomedical Engineering of the Medical University of Vienna (MUW). The system employs a 1340 nm central wavelength swept source with a maximum bandwidth of 27 nm and a repetition rate of 200 kHz. The lateral and axial resolutions were experimentally determined by imaging a resolution target and a mirror, to be 45.67 μm and 26.86 μm in air, respectively. The Field of View (FOV) of the images were 1cm $\times$ 1cm with a pixel size of 19 μm in the air (lateral direction) and 4.58 μm in the axial direction. The number of B-scans per position were four. The acquisition pattern consisted of acquiring consecutively 512 A-lines per B-scan (repeated 4 times) and movement to the next position. The output volume size of the samples were 1100 (pixel per A-line) $\times$ 512 (A-line per B-scan) $\times$ 2048 (512 B-scans $\times$ 4 times). For a further description of the setup and the system consult (Z. Chen et al. 2016).

4.2.3 Data

First, as a requirement of the method, we acquired a homogeneous scattering phantom made with synthetic clay (Blu-Tak ®Bostik, Wauwatosa, Wisconsin). We acquired 2048 B-scans in 512 different positions with the SS OCT of MUW with dimensions of 1100 $\times$ 512. With these images we computed the variance of the detail coefficients of 4 wavelet sub-bands. For the quantitative evaluation of the method we used 4 different datasets using the previously described OCT imaging systems. In the following table we describe the sample area and a description of the pathology or skin structure in each dataset.

Table 4.2. Datasets used in the assessment 256 OCT images in 64 consecutive positions with 4 B-scans per position to allow compounding. The initial dimensions of all images were 1100 $\times$ 512.

dataset / Id	Sample Area	Sex	Age	Description
1	Hand	Female	25	Nevus araneus
2	Back	Male	30	Recovery after surgery of a dysplastic naevus
3	Thigh	Female	27	Normal skin
4	Arm	Male	69	Basal Cell Carcinoma

For the preparation of the raw datasets we rescaled the volumes considering the axial (4.58 μm) and lateral resolution (19 μm). Figure 4.2 shows an example of a raw frame of the dataset 1 before (a) and after (b) scaling and of a frame of the phantom (c). For the calculation of the noise variance we consider the Region of Interest (ROI) marked with a white rectangle.

Figure 4.2. Examples of raw frames used in the study. (a) Frame of dataset 1 before rescaling considering axial and lateral resolution. (b) Frame after rescaling (c) Frame and ROI of the phantom used to calculate the noise variance.

4.2.4 Quantitative Evaluation

We evaluated the algorithm efficacy with common speckle-reduction performance metrics (Ozcan et al. 2007; Adler, Ko, and Fujimoto 2004; Zaki et al. 2017) and compare it with different state-of-art methods. We included as quantitative measurement the signal-to-noise ratio (SNR), the contrast-to-noise ratio (CNR) and the equivalent number of looks (ENL).

$$SNR = 10 \log_{10} \left(\frac{\mu_I^2}{\sigma_I^2} \right) \quad (8)$$

$$CNR = \left(\frac{1}{R} \right) \sum_{r=1}^R (\mu_r - \mu_b) / \sqrt{\sigma_r^2 + \sigma_b^2} \quad (9)$$

$$ENL = \left(\frac{1}{H} \right) \sum_{h=1}^H (\mu_h^2 / \sigma_h^2) \quad (10)$$

$$\beta = \frac{\Gamma(I_D - \mu_{I_D}, I_O - \mu_{I_O})}{\sqrt{\Gamma(I_D - \mu_{I_D}, I_D - \mu_{I_D}) \cdot \Gamma(I_O - \mu_{I_O}, I_O - \mu_{I_O})}} \quad (11)$$

SNR is defined in (8) where μ_I indicates the mean of the OCT image, and σ_I^2 refers to the noise variance. In the definition of CNR (9) μ_b and σ_b^2 are the mean and variance in a background noise region. μ_r , σ_r^2 are the mean and variance of all regions of interest (R), including the homogeneous and heterogeneous regions of interest. ENL is a measure of the smoothness of a homogeneous region of interest. In (10) μ_h , σ_h^2 are the mean and variance of all homogeneous regions of interest (H). The ratio β is used to evaluate the performance of the methods to preserve the structural features of the image. In (11) I_D refers to the

denoised images, I_O refers the original image, μ_{I_D} is the mean of the denoised image and μ_{I_O} is the mean of the original image. The operator Γ por two images I_1 and I_2 is defined as $\Gamma(I_1, I_2) = \sum_{i,j} [I_1(i,j)I_2(i,j)]$ with i and j refers to pixel position in the image. Except for the SNR calculations, all the other parameters were computed from the logarithmic OCT images.

To assess the influence of all the quantitative metrics we consider an additional ratio named figure-of-merit (FOM) (Ozcan et al. 2007) .

$$FOM = SNR_n + CNR_n + ENL_n + \beta_n \quad (12)$$

where n refers to the fact that the metric is normalized, i.e. the filter that performed the best in, for example, the SNR criteria, would have an SNR_n equal to one. Therefore FOM of 4 indicates that the filter performed the best in all the image quality ratios (SNR, CNR, ENL and β).

We compared the performance of our new WCAN algorithm with other state-of-art methods following two different approaches. First we considered as basic reference the average of the 4 frames (MEAN). We also compared with 2D filters that demonstrated good performance in speckle reduction such as the Hybrid Median Filter (HMF) (T. Huang, Yang, and Tang 1979) and the Bayesian Nonlocal means algorithm (OBNLM) (Coupe et al. 2009). Additionally we used another wavelet compounding filter, the Wavelet Multi-frame algorithm (WVMF) (Mayer et al. 2012). In the next figure we present the process used to evaluate all the methods. All the steps were performed for the 256 frames of every dataset in groups of 4 frames. The output denoised OCT frames of all the methods were compared using the performance metrics described (SNR, CNR, ENL and β).

Figure 4.3. Process used for the evaluation of the methods. The noise variance estimation is performed once for the reference stack ($RF_1, RF_2, \dots, RF_{2048}$). The compounding of the 4 initial OCT frames was done using WCAN, WVMF and MEAN. An additional digital filtering was performed from the averaged MEAN output using HMF and OBNML filters.

4.2.5 Qualitative Assessment

In addition to the proposition and quantitative evaluation of a new speckle denoising filter the other main objective of our study was to perform an initial assessment of the clinical value of the new method together with other state-of-the-art techniques. Although the quantitative ratio improvement is usually related with a better quality of the image and a potential benefit in the clinical evaluation and identification of signs that help in the diagnostic procedure, few studies considered a visual assessment with respect to clinical structures apart from a raw visual inspection of the denoised images. For this purpose we selected several criteria (table 4.3) relevant to the clinical evaluation of OCT skin images. Apart from global assessment fields (the image quality and the noise reduction ability) the experts had to evaluate the ability of the denoised method to enhance the typical structures present in the OCT skin images; the epidermis, the dermis, skin appendages, vasculature and pathologic structures. Several studies (Gambichler, Orlikov, et al. 2007; Olmedo et al. 2006; Hussain, Themstrup, and Jemec 2015; Mogensen et al. 2009; Gambichler, Pljakic, and Schmitz 2015; Forsea et al. 2010; Schmitz et al. 2013) reported the capacity of the OCT to visualize altered skin architecture and its correlation with pathologies like BC, AK, SCC, wound healing or the bowen disease among others. So the ability of the denoising techniques to enhance the study of features like the epidermal thickness, the disruption of layering,

dilated vessels, white streaks and dots in the epidermis or dark rounded areas, that could be present in the case of nonmelanoma skin cancer (Gambichler, Pljakic, and Schmitz 2015), is relevant for the daily application of OCT in dermatological practice. In the next figure we can appreciate examples of OCT skin imaging with significant structures marked. The stratum corneum (A.1) appears as a thin hyperreflective band at the top of the image corresponding to the surface of the skin. The epidermis (A.2) usually appears as a dark and homogenous layer with a well-defined border towards the upper dermis. The change in contrast seen between the epidermis and dermis is visualized as a fine hyporeflective line corresponding with the dermo-epidermal junction (DEJ) evident in histology. Skin appendages appear as hyporeflective discontinuations of the epidermal band. The dermis (A.3) is visualized as more hyperreflective with signal-poor cavities corresponding to appendages and vessels (A.5), and sometimes shadows cast from hairs (A.4) on underlying features. BCC lesions in OCT images may include alterations like hyporeflective rounded areas (B) in the dermis.

Figure 4.4. Example of OCT skin images with typical structures. The images were acquired with OCT device described in section 4.2.2 (A) Healthy OCT image of a 25 years old female. (1) Stratum corneum (2) Epidermis (3) Dermis (4) Hair follicle (5) Blood vessel. (B) BCC OCT image example of a 69 years old male. Red arrows indicate a typical pathological finding.

One specialist with more than 3 years of expertise in the use of OCT in skin imaging, assessed a pre-selected frame from the different datasets processed using the methods WCAN, WVMF, HMF, OBNLM and MEAN. The raw version of the datasets was also provided to the specialist to assess improvement. The order of the denoised and raw OCT images were randomized for each dataset to ensure a completely blinded procedure. The preselection of the frame to be assessed from each dataset was performed by the same specialist considering its clinical value and the capacity to evaluate the criteria indicated in table 4.3. The assessment was performed using the tele-screening tool described in chapter 3. The tool required the implementation of new views with the fields to be considered in the assessment.

Table 4.3. Criteria considered in the clinical assessment of the OCT denoised images. All the fields were graded from 1 to 4 (being 4 the best grade) in all the images denoised using all filters under study.

Criteria	Description
Epidermis	Identification of the epidermis layer and its borders.
Dermis	Identification of the dermis layer and its borders.
Skin adnexa	Ability to visualize hair follicle, glands, nails...
Vasculature	Identification of vasculatures and its thickness.
Noise reduction	Global evaluation of the noise reduction efficacy of the technique.
Image Quality	Global evaluation of the denoised image quality.

In the next figure we can see an example of the main view used in the evaluation. For each criteria used in the evaluation the expert could select from the values (Excellent, Good, Enough, Bad, Not evaluable). To be able to quantify the evaluation we give a numerical value to each option: 4 (excellent), 3 (good), 2 (enough) and 1(bad).

Figure 4.5. Example of OCT view of the tele-screening tool used by the evaluator. At the same view the expert had access to all the OCT images to evaluate and the fields that had to grade. Additionally the evaluator could include a global clinical evaluation of the image, although it was not used for this study.

UPM DermaTool - Quality Assessment Study

Images / Test (Juanjo Gómez)

1- OCT 2- OCT 3- OCT 4- OCT 5- OCT 6- OCT 7- OCT

Date: 21-03-2018 Description: Evaluate all the fields on the right of the image.

Download

☆ Epidermis:

☆ Dermis:

☆ Skin Adnexa:

Excellent
Good
Enough
Bad
Not evaluable

☆ Noise reduction:

☆ Global Quality:

Observations:

4.3 Results and Discussion

4.3.1 Quantitative evaluation

Before applying the method to all the datasets and compare to the rest of the filters we adjusted the parameters of the method. We used the dataset 1 for that purpose. We tested different number of decomposition levels (L), different values of the parameter k , used to balance the amount of noise reduction, and the number of B-scans to compound (N). The final values $L=4$, $k=1.5$ and $N=4$ were determined as an adequate balanced between performance ratio improvements and execution time.

Next, we present the results of the quantitative evaluation of the datasets described after the application of the selected methods for the complete stack of frames (256 frames for each dataset). We used the same ROIs for all the frames in the same dataset in order to maintain a fixed reference in the evaluation. The initial values of the metrics for all the datasets are included in table 4.4. From table 4.5 we can conclude that all the denoising techniques improve the image quality metrics (SNR, CNR and ENL). The best results are accomplished by the methods using the wavelet compounding strategies, WCAN and WVMF algorithms, with respect to the application of a mean compounding plus digital filtering (HMF, OBNLM). With this strategy the enhancement metrics increase in all the techniques and reduce the speckle noise, improving the possible study of details in the image (figure 4.7). The beta parameter determines the degree of alteration of the image in relation with the original image, as expected, the methods with more reduction of noise (WCAN and WVMF) presented a lower value of the beta parameter. The proposed method presents the best performance in CNR and ENL and the second in SNR (with WVMF being the first). The potential limitation in the edge preservation because of the lower value of the beta parameter with respect to the rest algorithms did not compromise the edge preservation according with the results of the qualitative evaluation that we will present in section 4.3.2.

Table 4.4. Median and standard deviation of the initial metrics value (raw data) for all the datasets included in the study.

dataset	SNR (dB)	CNR	ENL
1	13.47 ± 0.11	2.26 ± 0.05	25.31 ± 2.31
2	10.28 ± 0.48	2.32 ± 0.09	22.28 ± 2.08
3	14.09 ± 0.29	2.01 ± 0.13	74.86 ± 11.06
4	8.86 ± 0.13	1.69 ± 0.03	12.34 ± 0.95

Table 4.5. Quantitative evaluation considering the results of all datasets used in the study (Table 1) considering 4 B-scans per position (median $\pm$ standard deviation) of the improvement with respect to the initial metrics presented in Table 4.4 (raw images), except for the beta parameter that directly compares the preservation of structures with respect to the original image). The results are ordered by the FOM rate as defined in (12). Values corresponding with the best performance in each metric and the FOM are highlighted in bold.

Filter	$\Delta\text{SNR (dB)}$	ΔCNR	ΔENL	Beta	FOM
WCAN	5.86 ± 1.50	0.85 ± 0.25	26.60 ± 29.50	0.842 ± 0.039	3.81
WVMF	6.52 ± 1.12	0.81 ± 0.23	20.77 ± 22.08	0.868 ± 0.036	3.67
HMF	4.28 ± 0.94	0.59 ± 0.17	17.42 ± 19.00	0.893 ± 0.030	2.97
OBNLM	3.49 ± 2.24	0.44 ± 0.14	7.84 ± 14.83	0.917 ± 0.028	2.34
MEAN	2.39 ± 0.91	0.15 ± 0.06	1.94 ± 3.09	0.925 ± 0.026	1.61

If we analyze the results of the four datasets separately we found the same order of improvement in all the metrics as in the global results. WVMF always shows the best SNR enhancement with WCAN second, the proposed WCAN method presents the highest improvement results for CNR and ENL and the MEAN is the one with the highest beta value. In the next figure we present the value distribution of the metrics in the quantitative evaluation.

Figure 4.6. Value distribution of the ratios in the quantitative evaluation of all datasets used in the study (table 4.2) using 4 B-scans per position. The proposed method correspond with the forth column (WCAN) labeled in red.

Finally, we computed the CPU time of the algorithms in order to test the efficiency of the implementations used in the study. As we can see in table 6 the application of the WCAN method requires considerably less CPU time than the other wavelet alternative WVMF and is only overcome by the MEAN and the HMF filter.

Table 4.6. CPU Time calculated for the first 4 B-scans of dataset 1. WCAN stands for the proposed method without pre-processing. HMF and OBNLM represent the result applying both filters to output of the mean of the four B-scans.

Filter	CPU Time (seconds)
WCAN	2.8906
WVMF	32.8594
HMF	0.5781
OBNLM	3.9375
MEAN	0.3125

In the next figure we can appreciate the difference appearance of the OCT images after applying all previous methods.

Figure 4.7. Comparison of the application of speckle reduction methods in a 512 pixels $\times$ 299 pixels frame of the dataset of the imaging device of MUW with ID 1 using 4 B-scans at the same position. (a) OCT raw image and ROIs used for the calculation of the quality metrics. White rectangular region is used for noise estimation, red rectangles represent the homogeneous regions (H=4) and green rectangles the non-homogeneous regions. The sum of both are used to calculate the CNR (R=9). (b) OCT raw image with the detailed ROI and the scale. The rest of the images represent the filtered OCT image after applying the method (c) WCAN proposed method (d) HMF (e) WVMF (f) OBNLM (g) MEAN. The vertical white line corresponds to A-Line #130. (h)

4.3.2 Qualitative Evaluation

The quantitative assessment of the methods, based on the comparison of the improvement of values of a selection of visual criteria, like the ones presented in this study, offers a standard and consistent view of the potential benefit in the quality enhancement of the images and the reduction of speckle noise for a better clinical interpretation. To explore the real impact of this type of algorithms a specialist in skin OCT imaging interpretation performed an assessment of the concepts presented in the table 4.3 in all the denoised images. We summarize the main results of this evaluation in table 4.7. As we described in section 4.2.5 the evaluation values were 4(excellent), 3(good), 2 (enough) and 1(bad). The proposed WCAN method presented the best evaluation in all the criteria, except in the case of the epidermis identification. One remarkable result is the relatively low rates marked in the majority of the processed frames (with values between 2 and 4). In some cases like “skin adnexa“, none of the techniques offered an acceptable behavior according to the expert feedback. A detailed analysis of each dataset shows that the proposed WCAN method was always in the best position in the noise reduction and global quality metrics, but not for the epidermis and dermis identification metrics, where the best options was OBNLM. We will use this valuable feedback to improve the algorithm. The figure 4.8 shows examples of images presented to the evaluator.

Table 4.7. Qualitative evaluation (mean $\pm$ standard deviation) of all datasets used in the study (table 4.2). The criteria “vessels” and “skin adnexa” were evaluated only in the datasets 1 and 3. All the other criteria were evaluated in all the datasets.

Filter	Epidermis	Dermis	Skin adnexa	Vasculature	Noise reduction	Global quality
WCAN	2.25 $\pm$ 0.5	2.5 $\pm$ 0.58	2 $\pm$ 0	2.5 $\pm$ 0.71	3 $\pm$ 0	3 $\pm$ 0
WVMF	2.5 $\pm$ 0.58	2 $\pm$ 0.81	1.7 $\pm$ 0.71	1.5 $\pm$ 0.71	2.5 $\pm$ 0.58	2 $\pm$ 0.81
HMF	2.25 $\pm$ 0.96	2.25 $\pm$ 0.96	1.5 $\pm$ 0.71	1.5 $\pm$ 0.71	2 $\pm$ 0.58	1.5 $\pm$ 0.58
OBNLM	3.25 $\pm$ 0.5	2.5 $\pm$ 0.58	1.5 $\pm$ 0.71	1.5 $\pm$ 0.71	2.25 $\pm$ 0.5	2 $\pm$ 0
MEAN	2.5 $\pm$ 0.58	2.25 $\pm$ 0.96	1.5 $\pm$ 0.71	2 $\pm$ 1.41	1.75 $\pm$ 0.96	2 $\pm$ 0.81

Figure 4.8. Example of selected images evaluated by the specialist. (a, c, e) OCT raw images from dataset 1, 2 and 4 with detailed ROIs and scale. (b, d, f) OCT image after applying the WCAN method over the corresponding images together with the detailed ROI.

Further work and experiments need to be done to confirm these conclusions and determine the influence in the evaluation of the implementations used, the OCT setting, the datasets selected and the evaluator. The proposed methodology could be a valuable starting point to identify the areas of improvement of new proposed algorithms and a relevant indicator of the impact of these methods in the clinical field.

4.4 Conclusion

We have described a new method for speckle reduction of skin OCT volumes. The method is based in a multi-scale noise adaptive wavelet compounding strategy. The method requires the estimation of the noise variance in the wavelet domain of the OCT setting. In this study we used for this purpose a homogeneous scattering sample. The noise variance is used to compute the weights for the detail coefficients of each sub-band of every frame in the volume before compounding all the filtered frames. We applied the new method to four different OCT volumes (datasets) with 256 frames each, and we compared its results with state of art algorithms. The results of the quantitative evaluation based on four different metrics demonstrated, that the proposed method has the best performance and can suppress noise preserving structural information. Additionally, we propose a new methodology to perform a qualitative evaluation by OCT skin clinical specialist defining significant clinical criteria. We applied the previous scheme to several images selected based on their clinical relevance using the same methods used in the quantitative evaluation. The qualitative comparison confirms the good performance of the proposed method and suggests a promising potential application for OCT denoising to improve diagnosis based on visual assessment or as a previous step to segmentation or quantification image processing schemes.

Chapter 5

CAD system for glaucoma disease

Glaucoma detection in color fundus images is a challenging task that requires expertise and years of practice. In this study we exploit the application of different Convolutional Neural Networks (CNNs) showing the influence in the performance of relevant factors like the dataset size, the architecture and the transfer learning with fine tuning. We also compare the performance of CNNs with human evaluators and explore the influence of the integration of images and data collected from the clinical history of the patients. We achieved with VGG19 and fine tuning an AUC of 0.94 with sensitivity and specificity ratios similar to the expert evaluators of the study. The experimental results using three different datasets with 2313 images indicate that this solution can be a valuable option for prescreening computer aid system for the detection of glaucoma in large-scale screening programs.

This study was partly implemented with the collaboration of the Diagnostic Image Analysis Group of the Radboud University Medical Center in Nijmegen (The Netherlands) with the supervision of Professor Clarisa Sanchez and the assistance of the researchers Bart Liefers and Freerk Venhuizen. We thank their valuable support in the realization of this study.

5.1 Introduction

Glaucoma is a group of optic neuropathies characterized by progressive degeneration of retinal ganglion cells (Weinreb, Aung, and Medeiros 2014). Glaucoma is the leading cause of global irreversible vision loss with a prevalence for population aged 40-80 estimated in 3-4% (Tham et al. 2014). The number of people with glaucoma worldwide was estimated in 64.3 million in 2013, increasing to 76.0 million in 2020 and 111.8 million in 2040 (Tham et al. 2014). Because glaucoma is an asymptomatic condition until a relatively late stage the diagnosis is frequently delayed. Population-level surveys suggest that only 10-50% of people with glaucoma are aware they suffer the disease (Weinreb, Aung, and Medeiros 2014). As early diagnosis of the condition and treatment can prevent vision loss, glaucoma screening has been tested in numerous studies worldwide (Tielsch et al. 1991; C. Fleming et al. 2005; Maul and Jampel 2010; D. Zhao et al. 2017). Current studies show that glaucoma screening can be cost-effective in risk population (family history, black ethnicity, age) and can be improved using a test with initial automated classification followed by the assessment of a specialist (Burr et al. 2007).

The standard of care of a glaucoma screening consists on routine optometrist visits every 2-3 years, refereeing to an ophthalmologist in case of suspect of glaucoma to further tests and final confirmation of the diagnosis. A glaucoma screening study usually includes detailed medical history, slit lamp examination, visual field, fundus photography and a tonometry(Einarson et al. 2006; Ervin et al. 2012), other studies have evaluated the used of tests like scanning laser tomography (HRT) (Healey et al. 2010), optical coherence tomography (OCT) (G. Li et al. 2010) and scanning laser polarimetry with variable corneal compensation (GDx-VCC) (Yamada et al. 2000). The use of telemedicine as a glaucoma screening tool, commonly referred as teleglaucoma(Kassam et al. 2013), can accurately detect positive glaucoma cases, according to some studies, even better than in-person examinations (Thomas et al. 2014). Color Fundus images have been broadly used in telemedicine campaigns as diabetes retinopathy (Ciulla, Amador, and Zinman 2003) or age macular degeneration (AMD) (Bats et al. 2014) and also in glaucoma (K. Yogesan et al. 1999; Ahmed et al. 2013; Gupta, Sinha, and Dagar 2013).

Color fundus imaging has been often used in combination of image processing algorithms to aid in the detection and grading of eye diseases (Burlina et al. 2017; M. Niemeijer et al. 2010). The availability of digital fundus cameras in primary care settings and their extensive use in eye screening programs explain the interest in the diagnosis of glaucoma using this image modality. But the manual detection of glaucoma in color fundus images is a challenging task that requires specific expertise and years of practice. To overcome this situations great effort has been made on automatic detection of glaucoma at an early stage through image processing of color fundus images. We can distinguish four main changes in the retinal structures associated with glaucoma: optic nerve head variance, neuroretinal rim loss, retinal nerve fibre layer defects and peripapillary atrophy (Haleem et al. 2013). In Fig

5.1 some findings typically used to detect glaucoma in color fundus images by the ophthalmologist are represented. For a detailed explanation of the process of diagnosis of glaucoma review the section 3.2. The diagnosis based on color fundus images combines the assessment of both local and regional structural patterns of the optical disc region.

Figure 5.1. Example of findings use to detect glaucoma in color fundus images (Haleem et al. 2013). (a) Quantification of the optic cup to disc ratio (CDR). The reduction of the optic nerve fibres (potentially related with glaucoma) could provoke that the optic becomes larger with respect to the optic disc (b) The neuroretinal rim usually follows a pattern (ISNT rule) where the inferior region is broader than the superior, broader than the nasal, and broader than the temporal region. The alteration in this pattern is a sign of suspect of glaucoma.

To aid in the detection of glaucoma numerous image processing algorithms have been proposed. We can find works that focus on the localization and segmentation of the optic disc (M. D. Abramoff et al. 2007; Walter and Klein 2001; Mookiah, Acharya, Chua, Min, et al. 2013) and numerous detection of glaucoma algorithms based on the extraction of features from the image or transformed versions of the image to train different types of classifiers. The extracted features could explore relevant information present in the images, and have a better representation capacity than clinical measurements. Among others we can mention glaucoma detection based on a probabilistic combination of previously compressed features extracted from the pixel intensity values, the Fourier Transform (FT) and B-splines coefficients (Bock et al. 2010), or using higher order spectra analysis and texture-based features extracted from preprocessed images and a Support Vector Machines (SVM) classifier (U. R. Acharya et al. 2011), or with a feature extraction based on higher order spectra and discrete wavelet transform and a SVM classifier (Krishnan and Faust 2013; Mookiah et al. 2012), or using an empirical wavelet transform with a least-squares SVM (Maheshwari, Pachori, and Acharya 2017) or with an adaptive histogram equalization convolved with several filter banks processed to create local configuration patterns that feed a k-nearest neighbor (kNN) classifier (U. Rajendra Acharya et al. 2017). We can find other

implementations in (Dutta et al. 2014; Noronha et al. 2014; Muramatsu et al. 2009; Khan et al. 2013). The previous methods apply the approach of identifying features in the image in order to train a classifier with all the findings extracted directly from the image or from a transformed version of it (using wavelets, FT, high order spectra analysis...). At the end the different algorithms explore different aspects and transformations of the ONH to determine patterns representative of glaucoma. In Table 5.1 we summarize the datasets and performance of some of these methods. In this work we will apply a different approach to the glaucoma detection problem through the use of Convolutional Neural Networks (CNNs).

Convolutional networks, commonly known as one of the most popular deep learning algorithms for image analysis, have become very rapidly a very successful alternative for analyzing medical images. These methods could be considered as the evolution of the supervised techniques started at the end of the 1990s, where training data sets of previously classified images are used to develop the system. This strategy supersedes the previous approach based on feature extraction and posterior classification mentioned in the previous paragraphs. The new deep learning paradigm implies that computers can perform the feature learning and classification simultaneously. We can usually find in a deep learning algorithm a network (model) formed by many layers that transform an input data (images normally) to outputs (e.g. pathology present/absent). The most successful type of models for medical image analysis is a sub-class of neural networks called convolutional neural networks (CNN) that was introduced in the 1980s (Fukushima and Miyake 1982).

(Litjens et al. 2017) provided a thorough review of the current use of these techniques in medical analysis. The study mentions state-of-art applications of deep learning technology in the main topics of biomedical image processing: classification, object detection, segmentation or registration among others. (Shin et al. 2016) mentioned three mayor strategies that used CNNs to medical image classification problems: training from scratch, using off-the shell pre-trained CNN features, and conducting unsupervised CNN pre-training with supervised fine-tuning. Another method successfully applied in recent years and that we will use in this work is the transfer learning of CNN models pre-trained from natural image datasets and fine-tuned to solve medical imaging tasks.

The CAD systems and the screening programs have also progressively adopted the new approach. Several studies outperform state-of-art CAD systems on the mediastinal lymph node detection and interstitial lung classification (Shin et al. 2016), mammographic lesions (Kooi et al. 2017), lung nodules (J.-Z. Cheng et al. 2016) or skin lesions (Esteva et al. 2017). The image modality that we will use in this work, color fundus imaging, has also been successfully analyzed with CNNs achieving state-of-art performance or outperforming previous implementations. We can find since 2015 numerous studies applying CNN in retinal vessel segmentation (Huazhu Fu et al. 2016), image quality assessment (Mahapatra et al. 2016), segmentation of the optic disc and the optic disc cup (Zilly, Buhmann, and Mahapatra 2017; Sevastopolsky 2017), diabetic retinopathy detection (Michael David Abramoff et al.

2016), age-macular degeneration detection (Burlina et al. 2017) or hemorrhage detection (M. J. J. P. van Grinsven et al. 2016) among others. The CNN architectures successfully exploit both local and global features present in the images, being a proper tool for the detection of Glaucoma. Several studies have already tackle this problem in color fundus images using CNN. In (X. Chen et al. 2015) applied a six layers architecture to optic disc patches previously segmented. (Al-Bander et al. 2017) used CNN to extract features and train a SVM classifier to detect glaucoma images. Recently (H. Fu et al. 2018) presented a novel ensemble network based the application of different CNNs to the global fundus image and to different versions of optic disc region. Finally, in OCT there are also recent studies applying CNNs for glaucoma detection (Muhammad et al. 2017) or the segmentation of layers (Gopinath, Rangrej, and Sivaswamy 2018; Fang et al. 2017). In the next table we present a summary of the datasets used and the results reported.

Table 5.1. Summary of methods for the detection of glaucoma in color fundus images. In the datasets column we indicate the number of normal cases (-) and glaucoma cases (+). For the performance we used the reported metric used in the study: AUC, accuracy (Acc) and specificity (Sp) and sensitivity (Sn).

Authors	Method	Datasets	Performance
(Bock et al. 2010)	Pixel values FFT coefficients, B-spline and probabilistic SVM	Private (336 -/239 +)	AUC – 87% Acc–80%
(Krishnan and Faust 2013)	HOS, TT, DWT with SVM	Private (40 training/20 testing)	Acc – 91.67 %
(Maheshwari, Pachori, and Acharya 2017)	2D EWT and LS-SVM	Private (30-/30+) RIM-ONE (255-/250+)	Acc – 98.33 %
(U. Rajendra Acharya et al. 2017)	Texton, LCP features and KNN	Private (143-/559+)	Acc – 95,7%
(X. Chen et al. 2015)	CNN (6 layers)	ORIGA (168+/482-) SCES (46+/1676-)	AUC (83.1%-88.7%)
(Al-Bander et al. 2017)	CNN 23 layers and SVM	RIM-ONE (200+/255-)	Acc – 88.2%, Sn - 85% and Sp - 89.8%
(H. Fu et al. 2018)	Ensemble of 4 CNNs	ORIGA (168+/482-) SCES (1636-/46+)	AUC-91.83%, Sn (84.78%), Sp (83.80%)

Our work is aimed at developing a prescreening computer aid system for the detection of glaucoma in large-scale screening programs. In this chapter we will exploit and discuss the application of different CNN architectures to the classification of glaucoma with fundus color images. We will describe the results in the performance of different architectures as well as some transfer learning schemes from pre-trained ImageNet CNN models. Finally, we will consider the potential benefit of integrating basic clinical data collected in the screening studies of the patients in the final classification.

The rest of the chapter is structured in several sections. In section 5.2 we describe the data sets, the network architectures, the training process and the performance metrics used in the study. In section 5.3 we report the results of the different experiments. Finally, section 5.4 contains the summary of the study with the main conclusions.

5.2 Materials and Methods

5.2.1 Study Datasets

We used 2313 retinal fundus images in this work coming from three different datasets: two publicly available RIM-ONE (Fumero et al. 2011) and DRISHTI-GS (Sivaswamy et al. 2014)) and one private from a screening campaign performed in Hospital de la Esperanza in Barcelona (Spain). We considered two categories, glaucoma positives and normal. Glaucoma positives include the images evaluated by a specialist as suspect of glaucoma or with glaucoma.

The Open Retinal Image Database for Optic Nerve Evaluation (RIM-ONE) is an ophthalmic image group of databases designed in order to be a reference for the design of optic nerve head segmentation algorithms and in development of computer-aided glaucoma diagnosis. The database was created with the collaboration of three Spanish hospitals: Hospital Universitario de Canarias, Hospital Clínico San Carlos and Hospital Universitario Miguel Servet. Our study used the three releases included by the authors until now. The image set was designed in collaboration of 4 glaucoma experts. The camera used to capture the images was a Nidek AFC-210 background camera with a 21.1-megapixel Canon EOS 5D Mark II body. All the images are centered at the optic disc.

The dataset DRISHTI-GS consist of 101 retinal fundus images for optic disc segmentation. All images were collected at Aravind Eye Hospital in Madurai (India). Glaucoma patient selection was done by clinical experts based on findings during examination. The retinal images come from Indians patients of 40-80 years old. The images were taken with the eyes dilated, centered on the optic disk, with a field of view of 30-degrees and of dimension 2996x1944 pixels and PNG uncompressed image format.

Finally the ESPERANZA dataset consisted of 1446 color fundus images with a field of view 45 degrees, centered on the macula and including the optic disc from patients with age ranging from 55 to 86 years old. The retinal images were provided from the glaucoma detection campaign performed at Hospital Esperanza in Barcelona (Spain) to 1006 different patients and described in detail in chapter 3. During the examination of the patient, during the screening campaign, several data from the clinical history were collected (like age, family history of glaucoma, personal record of glaucoma in both eyes and the personal record of glaucoma therapy in both eyes among others) and besides the color fundus images, other tests (like the measurement of the intraocular pressure (IOP) in both eyes, the visual acuity tests and some Optical Coherence Tomography acquisitions) were performed. Special care was taken in order to create a reference gold standard dataset. All the images had a double

complete glaucoma evaluation performed by six ophthalmologist experts and nine non-expert using the tele-screening tool described in chapter 3. The evaluator ophthalmologists with more than five years of experience were considered experts in this work. In case of disagreement between the two evaluations two glaucoma experts decided the final category by consensus. The ophthalmologists evaluated a proportional part of all the images inside its category (see figure 5.2). The assessment of the images includes the evaluation of the quality of the images in the categories of good, enough, bad or not evaluable and the clinical classification in the categories of suspect of glaucoma and glaucoma. The inclusion criteria of the images of the campaign to the final dataset used in this work, includes the images that were labeled by the evaluators with quality good or enough and in the case of glaucoma positive cases, we included the images classified as suspect of glaucoma or glaucoma. In figure 5.2 we describe the previous process.

Figure 5.2. Gold standard ESPERANZA dataset. The total number of images acquired were 2049 color fundus images from 1006 studies. 176 were excluded by not enough image quality. 427 of the rest were excluded because they were classified with other pathology different from the glaucoma or because they were classified as normal but the OCT test of the patient were classified with glaucoma. All the initial 1006 studies were evaluated two times by six experts and nine non-experts. Each evaluator graded a proportional part of the hall dataset, i.e., each expert around 168 studies and each non-expert around 112 studies. In case of disagreement between the evaluation of the same study, two glaucoma experts decided by consensus. Consult chapter 3 for a detail description of the screening campaign.

Table 5.2 contains information of the retinal images considered in the study from each dataset. The initial size of the images of the RIM-ONE datasets are very different. In the case of *r1* the image sizes vary from 316x342 to 831x869. In the version *r2* the sizes vary from 290x290 to 1375x1654. Figure 5.3 shows images from the ESPERANZA dataset included in the positive and negative category and the findings identified by the clinicians using the tele-screening tool.

Table 5.2. Datasets used in the study. The term “glaucoma” includes all retinal images in the datasets classified by a specialist as suspect of glaucoma or as suffering the disease in any stage (early, moderate or severe glaucoma).

dataset	Initial Size	Format	Glaucoma	Normal
ESPERANZA	1024x680	JPG	113	1333
RIM-ONE r1	Not fixed	BMP	40	118
RIM-ONE r2	Not fixed	JPG	200	255
RIM-ONE r3	2144x1424	JPG	71	82
DRISHTI-GS	2996x1944	PNG	70	31
TOTAL			494	1819

Figure 5.3. Examples of color fundus images from the ESPERANZA dataset. All images were labeled with good quality by the evaluators. (a1-2) Left eye and optic disc detail of a suspect of glaucoma image with the ISNT rule violated (b1-2) Right eye and optic disc detail of a suspect of glaucoma image with rim thinning. (c1-2) Left eye and optic disc detail classified as normal.

As we mentioned before, the ESPERANZA dataset had a gold standard classification created by consensus of the evaluation of two groups of ophthalmologists, the experts and non-experts. We used this information to create a reference of the performance of human evaluators and to compare it with the performance of the proposed Deep Learning Architectures. In the next table we present a reference of the specificity and sensitivity of the expert and non-expert evaluators with respect to the gold standard. For the calculation of these metrics we considered color fundus images with quality good or enough. The gold standard set included 1735 color fundus images classified as normal and 160 classified with glaucoma or suspect of glaucoma. Consult section 3.4 for more details about the design and implementation of the campaign.

Table 5.3. Specificity and sensitivity (as defined in the performance metrics section 1.3.5) reference in the classification of the experts and non-experts evaluators respect to the gold standard in the ESPERANZA dataset. The values were calculated considering all the images evaluated by the ophthalmologists during the campaign with quality good or enough.

Class	Experts	Non-experts
Specificity	0.8914	0.8607
Sensitivity	0.7662	0.5875

5.2.2 Preprocessing

In the preprocessing step, we processed the images from the different datasets to a common and standard format in order to train the networks in a homogeneous way. No correction of illumination or contrast enhancement was applied to the images.

We decided to use standard patches centered at the optic disk and the same size for all the datasets, because of the clinical interest for the classification of this region and also to reduce the computational costs in the training step. In the case of the images from the ESPERANZA and DRISHTI-GS datasets we had to localize the optic disk to center the image and scaled them to 256x256 and 224x224 pixels size in order to adapt them to the different networks we used in the study. For the localization of the optic disk we used the same approach previously proposed by (Sekhar, Al-Nuaimy, and Nandi 2008) with the application of morphological operations of the binary image previously corrected to limit the effect of the blood vessels and small exudates on the image. A region of interest was then obtained to crop the image with the optic disc in the center. The described method localized correctly the optic disc with an accuracy of 84.02% for the images for the ESPERANZA dataset and 95.05% for the DRISHTI-GS dataset. The images with the optic disk wrongly identified were manually segmented. In the case of the different releases of the RIM-ONE dataset, we only had to perform a scaling to the final sizes (256x256 and 224x224 depending of the CNN). Finally we subtracted the mean across every channel, to ensure that all data inputs have the same centered distribution.

5.2.3 CNNs used and Transfer learning

For the selection of the network we tested three different CNNs methods. The first one (STANDARD CNN) used the architecture presented in Figure 5.4. It consists of 15 convolutional, pooling, fully-connected and softmax layers and was designed following standard CNN principles. At the end of the network a softmax layer performs the final binary classification (normal and glaucoma positive). The final output of the classifier is a two-element vector. In the next figure and table we present the details of the design of the network.

Figure 5.4. Architecture of the STANDARD CNN used in this work.

Table 5.4. Detail of the layers and parameters defined in the standard CNN. The input of the network is an image of size 256x256x3 and the output are two scores of the two classes considered. In the parameters column K indicates the number of filters of the layer

Layer	Operation	Input size	Parameters	Layer	Operation	Input size	Parameters
1	Convolution	256x256	3x3, K=32	9	Max-pooling	57x57	2x2
2	Convolution	254x254	3x3, K=32	10	Convolution	28x28	3x3, K=32
3	Max-pooling	252x252	2x2	11	Convolution	26x26	3x3, K=32
4	Convolution	126x126	3x3, K=32	12	Max-pooling	24x24	2x2
5	Convolution	124x124	3x3, K=32	13	Convolution	12x12	3x3, K=32
6	Max-pooling	122x122	2x2	14	Convolution	10x10	3x3, K=32
7	Convolution	61x61	3x3, K=32	15	Fully-connected	8x8	Dropout, p=0.5
8	Convolution	59x59	3x3, K=32	16	Soft-Max	128x1	2 classes

Training a deep CNN from scratch (full training) present relevant limitations. It requires large amount of labeled data, which in fields like medical imaging could be extremely expensive, especially for images that present pathological findings relevant for the diagnosis. Besides, the training of a deep CNN usually requires extensive memory and computational resources and it could be a very time consuming task. Finally, the design of a CNN and the adjustment of the hyper-parameters of the network could be a challenging process that requires dealing with overfitting and other issues that can limit the success of the application of this technology. One alternative to overcome these problems is the use of transfer learning with fine tuning. Transfer learning is a method successfully used in machine learning and data mining for classification, regression and clustering problems. It is generally defined as

the capability of the system to utilize the knowledge learned in one domain of interest, to another that shares some common characteristics (Pan and Yang 2010). It has been successfully tested in several medical image analysis studies (Shin et al. 2016; Karri, Chakraborty, and Chatterjee 2017; Huynh, Li, and Giger 2016; Tajbakhsh et al. 2016) the use of state-of-art performance CNNs on the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) (Deng et al. 2009; Russakovsky et al. 2015) (with millions of labeled images from 1000 different classes). As we will confirm in this work, the use of these architectures offers two valuable benefits. First, the design and architectural improvements of these CNNs can warrant superior performance ratios even in a training from scratch scheme. Second, the use of pre-trained deep CNNs and the subsequent fin-tune of the weights of the network applying the new labeled images, could lead to even better performance metrics and a potential reduction in training resources in terms of time, memory and computational operations. Besides the STANDARD CNN presented in Fig 5.4, we selected two additional CNNs (VGG19 and GoogLeNet) and prepared several experiments to analyze quantitatively the contribution of the selection of the architecture, the fine-tuning (VGG19_TL, GOOGLELENET_TL) versus full training (VGG19, GOOGLELENET), and the dataset size.

VGG19 (Simonyan and Zisserman 2014) is a publicly available CNN model that includes five stacks, each stack contains between two and four convolutional layer followed by a max-pooling layer, and it ends with three fully connected layers (figure 5.5). The main contribution of this architecture was the increasing of the depth of the network (in this work we applied the version with 19 layers) and the use of very small (3x3) convolutional filters. We can find an example of the application of this network for the fixed features extraction in CADx for breast cancer detection with different imaging modalities (Antropova, Huynh, and Giger 2017).

Figure 5.5. VGG19 Architecture with filter dimensions and the number of filters per layer.

GoogLeNet proposed in (Szegedy et al. 2015) accomplished as main contribution the improved utilization of the computing resources inside the network, increasing the depth and width of the network but keeping the computational budget constant. It also proposed a new module called “Inception” which concatenates convolution layers with different kernel sizes and one pooling layer into a single new filter. The complete architecture (figure 5.6) contains 22 layers including two convolution layers, three pooling layers and nine inception layers.

GoogLeNet was successfully used in the detection of lymph node metastases in women with breast cancer (Bejnordi et al. 2017), in the classification of normal and cancerous lung tissues from CARS (Coherent anti-Stokes Raman scattering) images (Weng et al. 2017) or retinal pathologies using optical coherence tomography (OCT) images (Karri, Chakraborty, and Chatterjee 2017).

Figure 5.6. GoogLeNet Architecture with filter dimensions and illustration of inception layer.

The input in VGG19 and GoogleNet networks were the preprocessed color fundus images from the datasets of the study with a final size of 224x224x3 and centered at the optic disc. The input images in the case of the standard network were 256x256x3. The region of interest presented to all the networks were the same. We changed the last layer with the softmax classifier in VGG19 and GoogleNet to consider only the two classes of interest in our study (glaucoma positives and normal).

5.2.4 Performance metrics

We define sensitivity, or true positive rate, as the number of true positives (number of images with glaucoma correctly detected) divided by the sum of the number of true positives and false negatives (images incorrectly classified as normal). Therefore, the sensitivity shows the percentage of glaucoma cases correctly identified by the algorithm. We define specificity as the number of true negatives (number of images normal correctly detected in our case) divided by sum of the number of true negatives and the false positives (images incorrectly classified as glaucoma). The specificity is ratio that shows the percentage of normal cases correctly identified. We will use the receiver operating characteristics (ROC) graph for visualizing the performance of the networks (Fawcett 2006). The ROC graph is a two dimensional representation with the sensitivity in the Y axis and 1-specificity in the X axes. We will compare the performance of the algorithms using the area under the receiver operating curve (AUC) generated by ROC curve. For the calculation of the optimum threshold we will consider the Youden index, defined as the index where the sum of the specificity -1 and sensitivity is maximum (Fluss, Faraggi, and Reiser 2005).

5.2.5 Training and testing processes

The complete dataset was randomly split in three different groups: training, validation and test set with the distribution of images presented in table 5.5. The validation set was used to monitor the hyper-parameters of the training process. The hyper-parameters that presented better performance in terms of accuracy in both classes (glaucoma positives and normal) with the validation set are presented in the table 5.6. After the selection of the hyper-parameters and the cross validation experiments, the rest of the trainings in the study considered in the training set also the images validation set (in total 370 glaucoma positives and 1364 normal). The test sets were the same during all the experiments (124 glaucoma positives and 455 normal).

Table 5.5. Distribution of images in the groups “glaucoma positives” and “normal”.

Group	Training	Validation	Test
Glaucoma Positives	333	37	124
Normal	1227	137	455

The overfitting is a well-known issue in CNNs with limited training data. In order to limit this important problem in the training process we applied different strategies commonly used for this purpose. First, the selection of the number of epochs to complete the process was stopped when the performance on the validation set was reduced. Second, we applied dropout, the technique presented by Srivastava et.al.(Srivastava et al. 2014) that consists on temporarily removing units along with all its incoming and outgoing connections in a neural network. We included dropout with $p=0.5$ in the standard CNN and in VGG19. Finally, we applied data augmentation. This technique consists of training and/or testing on systematically transformed images. The transformations used typically have to maintain the

classification of the original image. In our study we used data augmentation during training in all the networks. In every iteration, every image included in the batch could be transformed by a random combination of the operations: random flip, random small rotations between -10 degrees and +10 degrees and random translation of maximum 40 pixels in the x or y direction of the image.

5.2.6 Implementation

The preprocessing steps were implemented using Matlab R2016b 64-bits (Mathworks, Inc.) on a desktop computer equipped with an Intel Xeon CPU ES31245 and the CNN experiments were implemented in Python (version 2.7.12) using the libraries Lasagne (version 0.2) (Sander Dieleman et al. 2015) and Theano (version 0.9) (Bastien et al. 2012) in a desktop computer with a NVIDIA GeForce GTX Titan Pascal 12GB GPU.

5.3 Results and discussion

We defined several experiments to evaluate the best solution in terms of performance and to get more insight of the different alternatives tested.

5.3.1 Hyper-parameter selection

As described before we used the validation set to select the best hyper-parameters for all the networks and types of training (scratch/fine tuning) used during the study. In the next table we present the final parameters selected. It is important to mention the relevant difference between the fine tuning networks and those with full training.

Table 5.6. Hyper-parameters selected after the monitoring of the training process using the validation set.

Method	Number of epochs	Learning rate	Batch Size
Standard CNN	100	0.005	64
VGG19	100	0.0001	32
GoogLeNet	100	0.0001	32
VGG19 fine tuning	20	0.0001	32
GoogLeNet fine tuning	40	0.0001	32

5.3.2 CNN algorithm selection

After the selection of the hyper-parameters we evaluated all the architectures under study (Standard CNN, VGG19 and GoogLeNet) in terms of the performance metrics, using all the data included in the training and validation subsets for training, and for testing the full test set. In the next figure and table we present the ROC curves and the performance metrics. The two first options are VGG19 with fine tuning (VGG19_TL) and GoogLeNet (GOOGLENET) with full training although this last option needed 100 epochs with respect to 20 epochs in the case of VGG19_TL. These results demonstrate that CNNs can be trained, using large data sets and without having to specify lesion-based features, to identify

glaucoma in retinal fundus images with high sensitivity and high specificity. Besides the best option achieves state-of-art performance ratios with images coming from multiples sources and different sizes and formats.

Figure 5.7. Global performance comparison ROC curves and AUC values of the comparison of the networks of the study using the test set described in table 5.5. The two first options in terms of AUC are labeled as (1) and (2).

Table 5.7. Performance ratios of all the CNNs under study ordered by the AUC value. The sensitivity and specificity was calculated using the Youden index previously described. The best option is presented in bold.

Method	AUC	Sensitivity (%)	Specificity (%)
STANDARD CNN	0.8742	81.45	83.74
VGG19	0.9082	81.45	83.74
GOOGLNET	0.9246	87.90	85.05
VGG19_TL	0.9436	85.48	91.64
GOOGLNET_TL	0.8979	88.71	78.68

5.3.3 Ten-fold cross validation

To confirm robustness and stability of the VGG19_TL network, a 10 fold cross validation test was performed considering all the data included in the training and validation subsets to make the different splits of the cross validation. The results showed an acceptable variability in the performance of the network with a standard deviation of AUC of 0.02 in the Mean ROC. The next figure shows the corresponding 10-fold ROC curves and the median of all the cases.

Figure 5.8. ROC curves and AUC values for the 10 fold cross validation experiment

5.3.4 CNN/human evaluator performance

We compared the classification of the best network of previous experiments with the performance of the experts and non-experts evaluators. Like in (Gulshan et al. 2016) we selected two operating points. The first one related with the specificity and the second with the sensibility for detecting glaucoma in both groups of evaluators, the ophthalmologist experts and the non-experts. The results are presented in the next figure, where we can appreciate the good performance of the solution compared with both groups of evaluators. According with these results the proposed method can achieve high sensitivity and specificity, with ratios comparable to an expert ophthalmologist with more than 5 years of experience. Screening populations for a high prevalence disease like glaucoma, require both high sensitivity and high specificity to minimize both false-positive and false-negative results.

Figure 5.9. ROC curve of VGG19 with fine tuning (VGG19_TL) over the test set of table 5.5 and relevant operating points. (a) High specificity operating point of the ROC curve with sensitivity 83.95% and specificity 90.32% (b) High sensibility operating point of the ROC curve with sensitivity 79.34% and specificity 92.74% (c) Reference performance of and an ophthalmologist expert in glaucoma considering all the ESPERANZA data set, sensitivity 76.62% and specificity 89.14%. (d) Reference performance of an ophthalmologist non expert in glaucoma considering all the ESPERANZA data set, sensitivity 58.75% and specificity 86.07 %.

In figure 5.10 we show the different behavior of the network with the different datasets. We can notice the inferior performance in the ESPERANZA dataset compared to the other datasets (DRISHTI-GS and RIM-ONE). There are several reasons that could explain this behavior. First, although the ESPERANZA dataset had a double evaluation with consensus it is a screening labeled dataset and further tests are needed to confirm the diagnosis, having considered for this study the positive cases suspected of glaucoma by assessing the fundus images. Second the localization of the optic disc is not the conventional for this type of assessment. If we review the section 3.4 we can see that the color fundus images acquired were centered in the macula and that the optic disc was at the border of the image. In the other datasets all the images are centered at the optic disc. Finally the initial image size of the global image in the ESPERANZA dataset (1024x680) suppose a possible limitation compared with the images of the other dataset which are already segmented (RIM-ONE) or have much more resolution (DRISHTI-GS). Further test should be performed to confirm these hypothesis and increase the sensitivity of the ESPERANZA dataset.

Figure 5.10. Confusion matrixes for the test dataset. (a) All the test dataset of the study. (b) Test set from the ESPERANZA database. (c) Test set from DRISHTI-GS database. (d) Test set from RIM-ONE database.

The figure 5.11 shows examples of the classification of the network in the ESPERANZA dataset with correct and incorrect classifications. If we focus on the false negative examples (e and f), that, as we mentioned before, had inferior performance than the other datasets, it is remarkable that for example in the case of (e) the non-expert labeled the image as normal and the expert as glaucoma. Finally it was classified as glaucoma because the glaucoma experts identified a violation in the ISNT rule. The case of (f) is similar, the non-expert classified as normal and the expert as glaucoma, and the glaucoma experts labeled the image as being affected by rim thinning and with a ISNT rule violation. This is not the case of the false positives (g, h) where the expert and non-expert evaluators labeled the image as “normal” but the algorithm identified them as “glaucoma”

Figure 5.11. Example of the classification of the VGG19_TL network for images from the test set of the ESPERANZA database. (a,b) True negative examples, the human evaluators and the algorithm identified the images as normal. (c,d) True positives examples, the human evaluators and the algorithm identified the images as suspect of glaucoma positives. (e,f) False negatives examples. The human evaluation identified both images as suspect of glaucoma but the algorithm labeled them as normal. (g,h) False positives examples. The human evaluators marked the images as normal but the algorithm classified them as glaucoma suspects.

5.3.5 Datasets performance influence

We evaluated the improvement in the performance of VGG19_TL associated with the inclusion of the different datasets for training. For that purpose we evaluated the network trained with one dataset (ESPERANZA), two (ESPERANZA and DRISHTI-GS) and the final dataset adding RIM-ONE. To study the performance of each trained model we always used the full test set that includes data of the three datasets. The full test dataset was used to verify the performance in the three training configurations. In the figure 5.12 we can verify that the evolution in the performance is clear starting in 0.8505 with 113 glaucoma positives and 1333 normal cases to 0.9436 with the full dataset with 494 glaucoma positives and 1819 normal cases.

Figure 5.12. ROC curves and AUC values using always the test set of Table 5.5 and training VGG19 with fine tuning and 10 epochs. The red line represents the ROC curve using the network trained only with the EPSESPANZA dataset. The blue line is the result after training with ESPERANZA and DRISHTI-GS datasets and the green line is the ROC curve after training using all the datasets of the study: ESPERANZA, DRISHTI-GS and RIM-ONE.

5.3.6 Integration of medical data and CNN

The final experiment consisted in the integration of clinical history data collected during patient's screening visit with color fundus images. As we have mentioned, CNNs are especially successful analyzing medical images. But the clinical diagnosis usually involves the assessment of a variety of exams that includes medical images but also data coming from different sources, like the medical history of the patient or from tests where the output could be a simple value. From this perspective, considering only one test could limit the final diagnosis and represent only a part of the global disease complexity. In order to consider a broad view in the diagnosis, some studies have explored the application of CNN integrating different imaging modalities (Xu et al. 2016; Liu et al. 2015) but also combining images with raw data, like in (Mobadersany et al. 2018) where histological images and genomic data were integrated in a single CNN. Following this approach we designed one experiment with the ESPERANZA dataset integrating color fundus images with medical history data collected from the same patient during the screening campaign. The data selected from the exam were: the age, the intraocular pressure (IOP) in both eyes, the family history of glaucoma, the personal record of glaucoma in both eyes and the personal record of glaucoma related therapy in both eyes. The selection of the data was based on some of the major risk factors of glaucoma presented in the literature (Weinreb, Aung, and Medeiros 2014; A. L. Coleman and Miglior 2008). In the next tables we present more information of the clinical data used in the experiment. This data was collected during the examination of the patients in the ophthalmological screening campaign described in chapter 3, review section 3.4 for more information.

Table 5.8. Mean IOP and standard deviation and mean age and standard deviation in the training set and test set in the glaucoma and normal classification.

Data	Glaucoma Set	Normal set
	Mean $\pm$ standard deviation	Mean $\pm$ standard deviation
IOP training set (mmHg)	13.94 $\pm$ 3.16	13.97 $\pm$ 2.79
IOP test set (mmHg)	13.64 $\pm$ 2.79	14.20 $\pm$ 2.87
Age training set (years)	70.99 $\pm$ 8.56	68.81 $\pm$ 7.42
Age test set (years)	70.81 $\pm$ 8.51	68.34 $\pm$ 6.94

Table 5.9. Percentage of cases with glaucoma family history, personal record of glaucoma and personal related therapy in the training and test set of the glaucoma and normal classification.

Data	Glaucoma Set	Normal set
Family history of glaucoma training	8.75 %	11.76 %
Family history of glaucoma test	12.12%	11.83 %
Personal record of glaucoma training	16.25 %	2.61 %
Personal record of glaucoma test	12.12 %	2.66 %
Personal glaucoma related therapy training	0 %	0.80 %
Personal glaucoma related therapy test	0 %	0.59 %

Color fundus images and the examination data were used together during the training by including the raw data into the last fully connected layer of the network. All the cases from the training and validation subsets from ESPERANZA were used for training (80 glaucoma positives and 995 normal) and the ESPERANZA cases from the test dataset were used for testing (33 glaucoma positives and 338 normal). For this experiment we only used the ESPERANZA dataset, because is the only dataset that contains information of the examination of the patients. We trained the network for 100 epochs. We performed two different experiments: the first one using all previous data and the second only with the age and the personal record of glaucoma because according with table 5.8 and 5.9 are the only data which appears different in the “glaucoma” and “normal” groups. As we can see in the next table, in the AUC values no significant difference is appreciated by adding the clinical history data, but the sensitivity and specificity has higher values which indicate that this information could be valuable to improve the classification. Although the results are promising, further tests with other integration architectures and other clinical fields should be considered.

Table 5.10. AUC, specificity and sensitivity values corresponding with the result of applying the network trained with the ESPERANZA dataset considering only the color fundus images (reference), integrating the images with all the data collected from the examination of the patient (fusion all data) and integrating only the age and personal record of glaucoma (fusion age/personal record glaucoma)

Group	AUC	Sensitivity (%)	Specificity (%)
Reference	0.8505	81.45	74.51
Fusion all data	0.8405	81.82	80.18
Fusion age/personal record glaucoma	0.8443	84.85	78.99

5.4 Conclusions

In this chapter, we exploit and extensively evaluate the application of deep convolutional neural networks for glaucoma detection using color fundus images in the context of large screening campaigns. We studied the influence of the architecture, the dataset size, the training strategy and the integration of data collected from the clinical history and the patient examination. We used three different datasets, two publicly available DRISHTI-GS and RIM-ONE, and other created from a glaucoma screening campaign to assess the performance of the alternatives. The three architectures tested, standard CNN, VGG19 and GoogLeNet offered good performance ratios in terms of AUC, sensitivity and specificity. The best option for the dataset used was VGG19 with transfer learning and fine tuning, with an AUC of 0.94, a sensitivity of 85.48% and a specificity of 91.64%, which showed a similar performance with respect to the expert evaluators of the screening campaign. We confirmed the great influence of the number of images and datasets using CNNs. The AUC of the VGG19 with fine tuning increase from 0.8505 with one dataset to 0.9436 with all the datasets of the study. Finally, we evaluated the performance of the integration of the data from the clinical history and tonometry tests with the color fundus images. The results show a slight improvement in sensitivity and specificity although the AUC was similar. Further tests with more data and new architectural approaches needs to be done to confirm this line of work. The good results presented demonstrated that CNNs on fundus images are a valuable alternative as a CAD system in glaucoma detection campaigns.

Chapter 6

Contributions

The main contributions of this work are listed below:

- Creation of a multi-specialty / multi-disease store-and-forward tele-screening tool. The tool is a secured web-based application that covers all the tasks needed in a screening campaign, from the recruitment of the patients to the inclusions of the examination, the remote evaluation of all the studies, the generation of reports, the integration with Medical Record System of the hospital and the inclusion of the final revision of the patients in the Hospital. The platform offers real-time statistics of the evolution of a campaign and exportation facilities of all the data for a posterior analysis or use in CAD systems.
- Application of the tele-screening tool in an ophthalmology screening campaign with 1006 patients. The application was used successfully by more than 30 health professionals in a three years project and helped in the detection of 31 cases of diabetic retinopathy, 168 cases of AMD and 154 of glaucoma among other pathologies. The tool showed the relevance of this type of campaigns in high prevalence pathologies like AMD and glaucoma where the number of cases with a pathological evaluation was high compared with other pathologies like diabetic retinopathy where the monitoring of the diabetic patients showed inferior positives cases.
- Design of a new enhancement method for OCT skin images based on wavelet compounding of several frames acquired at the same position. The application of the method on four datasets accomplished state-of-art performance ratios showing their potential application in newly available portable OCT systems that could be used in screening campaigns, in CAD systems for improvement in the detection of skin diseases, or as pre-processing step in image processing tasks.
- Creation of a methodology for quality assessment of medical images in order to evaluate the clinical value of the application of image enhancement algorithms. The method was successfully tested with the enhancement method proposed for OCT images, using the tele-screening tool. To the best of our knowledge OCT enhancement methods have not been compared with a similar methodology before this work.

- Demonstration of the application of deep convolutional networks for the automatic detection of glaucoma. We studied the influence of the architecture, the dataset size, the training strategy and the integration of data collected from the clinical history and the examination of the patients.
- Creation of an automatic algorithm for the detection of suspects of glaucoma in screening campaigns using fundus image based on the VGG19 network and using transfer learning and fine tuning resulting in a performance of an AUC of 0.94, a sensitivity of 85.48% and a specificity of 91.64%, which showed a performance similar to the one achieved by the expert evaluators of the screening campaign, and clearly superior with respect to the non-expert evaluators. The solution could be a valuable tool to be incorporated in CAD systems for glaucoma detection campaigns.

References

- Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, et al. 2016. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems." *ArXiv:1603.04467 [Cs]*, March. <http://arxiv.org/abs/1603.04467>.
- Abràmoff, M. D., Wallace L. M. Alward, Emily C. Greenlee, Lesya Shuba, Chan Y. Kim, John H. Fingert, and Young H. Kwon. 2007. "Automated Segmentation of the Optic Disc from Stereo Color Photographs Using Physiologically Plausible Features." *Investigative Ophthalmology & Visual Science* 48 (4): 1665–73. <https://doi.org/10.1167/iovs.06-1081>.
- Abramoff, M. D., M. K. Garvin, and M. Sonka. 2010. "Retinal Imaging and Image Analysis." *IEEE Reviews in Biomedical Engineering* 3: 169–208. <https://doi.org/10.1109/RBME.2010.2084567>.
- Abràmoff, Michael David, Yiyue Lou, Ali Erginay, Warren Clarida, Ryan Amelon, James C. Folk, and Meindert Niemeijer. 2016. "Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep LearningDeep Learning Detection of Diabetic Retinopathy." *Investigative Ophthalmology & Visual Science* 57 (13): 5200–5206. <https://doi.org/10.1167/iovs.16-19964>.
- Acharya, U. R., S. Dua, X. Du, V. Sree S, and C. K. Chua. 2011. "Automated Diagnosis of Glaucoma Using Texture and Higher Order Spectra Features." *IEEE Transactions on Information Technology in Biomedicine* 15 (3): 449–55. <https://doi.org/10.1109/TITB.2011.2119322>.
- Acharya, U. Rajendra, Shreya Bhat, Joel E. W. Koh, Sulatha V. Bhandary, and Hojjat Adeli. 2017. "A Novel Algorithm to Detect Glaucoma Risk Using Texton and Local Configuration Pattern Features Extracted from Fundus Images." *Computers in Biology and Medicine* 88 (September): 72–83. <https://doi.org/10.1016/j.compbimed.2017.06.022>.
- Adler, Desmond C., Tony H. Ko, and James G. Fujimoto. 2004. "Speckle Reduction in Optical Coherence Tomography Images by Use of a Spatially Adaptive Wavelet Filter." *Optics Letters* 29 (24): 2878–80. <https://doi.org/10.1364/OL.29.002878>.
- Agurto, Carla, E. Simon Barriga, Victor Murray, Sheila Nemeth, Robert Crammer, Wendall Bauman, Gilberto Zamora, Marios S. Pattichis, and Peter Soliz. 2011. "Automatic Detection of Diabetic Retinopathy and Age-Related Macular Degeneration in Digital Fundus Images." *Investigative Ophthalmology & Visual Science* 52 (8): 5862–71. <https://doi.org/10.1167/iovs.10-7075>.
- Ahmed, Rezwan, Stephen Petrany, Russell Fry, and Michael Krasnow. 2013. "Screening Diabetic and Hypertensive Patients for Ocular Pathology Using Telemedicine Technology in Rural West Virginia: A Retrospective Chart Review." *Internal Medicine*, January. http://mds.marshall.edu/int_med/24.
- Al-Bander, B., W. Al-Nuaimy, M. A. Al-Tae, and Y. Zheng. 2017. "Automated Glaucoma Diagnosis Using Deep Learning Approach." In *2017 14th International Multi-Conference on Systems, Signals Devices (SSD)*, 207–10. <https://doi.org/10.1109/SSD.2017.8166974>.

- Andermann, Anne, Ingeborg Blancquaert, Sylvie Beauchamp, and Véronique Déry. 2008. "Revisiting Wilson and Jungner in the Genomic Age: A Review of Screening Criteria over the Past 40 Years." *World Health Organization. Bulletin of the World Health Organization*; Geneva 86 (4): 317–19.
- Antropova, Natalia, Benjamin Q. Huynh, and Maryellen L. Giger. 2017. "A Deep Feature Fusion Methodology for Breast Cancer Diagnosis Demonstrated on Three Imaging Modality Datasets." *Medical Physics* 44 (10): 5162–71. <https://doi.org/10.1002/mp.12453>.
- Apalla, Zoe, Dorothee Nashan, Richard B. Weller, and Xavier Castellsagué. 2017. "Skin Cancer: Epidemiology, Disease Burden, Pathophysiology, Diagnosis, and Therapeutic Approaches." *Dermatology and Therapy* 7 (1): 5–19. <https://doi.org/10.1007/s13555-016-0165-y>.
- Armstrong, William B., James M. Ridgway, David E. Vokes, Shuguang Guo, Jorge Perez, Ryan P. Jackson, Mai Gu, et al. 2006. "Optical Coherence Tomography of Laryngeal Cancer." *The Laryngoscope* 116 (7): 1107–13. <https://doi.org/10.1097/01.mlg.0000217539.27432.5a>.
- Arora, Sourabh, Chris J. Rudnisky, and Karim F. Damji. 2014. "Improved Access and Cycle Time with an 'In-House' Patient-Centered Teleglaucoma Program Versus Traditional In-Person Assessment." *Telemedicine and E-Health* 20 (5): 439–45. <https://doi.org/10.1089/tmj.2013.0241>.
- Bach, Peter B., Joshua N. Mirkin, Thomas K. Oliver, Christopher G. Azzoli, Don Berry, Otis W. Brawley, Tim Byers, et al. 2012. "Benefits and Harms of CT Screening for Lung Cancer: A Systematic Review." *JAMA: The Journal of the American Medical Association* 307 (22): 2418–29. <https://doi.org/10.1001/jama.2012.5521>.
- Bahaadinbeigy, Kambiz, and Kanagasingam Yogesan. 2012. "A Literature Review of Teleophthalmology Projects from Around the Globe." In *Digital Teleretinal Screening*, 3–10. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-25810-7_1.
- Bai, V. Thulasi, V. Murali, R. Kim, and S.k. Srivatsa. 2007. "Teleophthalmology-Based Rural Eye Care in India." *Telemedicine and E-Health* 13 (3): 313–21. <https://doi.org/10.1089/tmj.2006.0048>.
- Baltussen, Rob, Jeroen Naus, and Hans Limburg. 2009. "Cost-Effectiveness of Screening and Correcting Refractive Errors in School Children in Africa, Asia, America and Europe." *Health Policy* 89 (2): 201–15. <https://doi.org/10.1016/j.healthpol.2008.06.003>.
- Bashkansky, M., and J. Reintjes. 2000. "Statistics and Reduction of Speckle in Optical Coherence Tomography." *Optics Letters* 25 (8): 545–47. <https://doi.org/10.1364/OL.25.000545>.
- Bastien, Frédéric, Pascal Lamblin, Razvan Pascanu, James Bergstra, Ian Goodfellow, Arnaud Bergeron, Nicolas Bouchard, David Warde-Farley, and Yoshua Bengio. 2012. "Theano: New Features and Speed Improvements." *ArXiv:1211.5590 [Cs]*, November. <http://arxiv.org/abs/1211.5590>.
- Bats, F. De, C. Vannier Nitenberg, B. Fantino, P. Denis, and L. Kodjikian. 2014. "Age-Related Macular Degeneration Screening Using a Nonmydriatic Digital Color Fundus Camera and Telemedicine." *Ophthalmologica* 231 (3): 172–76. <https://doi.org/10.1159/000356695>.

- Baumeister, Thomas, Wobbeke Weistenhöfer, Hans Drexler, and Birgitta Kütting. 2009. "Prevention of Work-Related Skin Diseases: Teledermatology as an Alternative Approach in Occupational Screenings." *Contact Dermatitis* 61 (4): 224–30. <https://doi.org/10.1111/j.1600-0536.2009.01606.x>.
- Bejnordi, Babak Ehteshami, Mitko Veta, Paul Johannes van Diest, Bram van Ginneken, Nico Karssemeijer, Geert Litjens, Jeroen A. W. M. van der Laak, et al. 2017. "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer." *JAMA* 318 (22): 2199–2210. <https://doi.org/10.1001/jama.2017.14585>.
- Berg, Wendie A., Zheng Zhang, Daniel Lehrer, Roberta A. Jong, Etta D. Pisano, Richard G. Barr, Marcela Böhm-Vélez, et al. 2012. "Detection of Breast Cancer With Addition of Annual Screening Ultrasound or a Single Screening MRI to Mammography in Women With Elevated Breast Cancer Risk." *JAMA* 307 (13): 1394–1404. <https://doi.org/10.1001/jama.2012.388>.
- Bezerra, Hiram G., Marco A. Costa, Giulio Guagliumi, Andrew M. Rollins, and Daniel I. Simon. 2009. "Intracoronary Optical Coherence Tomography: A Comprehensive Review: Clinical and Research Applications." *JACC: Cardiovascular Interventions* 2 (11): 1035–46. <https://doi.org/10.1016/j.jcin.2009.06.019>.
- Bhan, Archana, Andrew C. Browning, Sunil Shah, Robin Hamilton, Dinesh Dave, and Harminder S. Dua. 2002. "Effect of Corneal Thickness on Intraocular Pressure Measurements with the Pneumotonometer, Goldmann Applanation Tonometer, and Tono-Pen." *Investigative Ophthalmology & Visual Science* 43 (5): 1389–92.
- Bibbins-Domingo, Kirsten, David C. Grossman, Susan J. Curry, Karina W. Davidson, Mark Ebell, John W. Epling, Francisco A. R. García, et al. 2016. "Screening for Skin Cancer: US Preventive Services Task Force Recommendation Statement." *JAMA* 316 (4): 429–35. <https://doi.org/10.1001/jama.2016.8465>.
- Bleyer, Archie, and H. Gilbert Welch. 2012. "Effect of Three Decades of Screening Mammography on Breast-Cancer Incidence." *New England Journal of Medicine* 367 (21): 1998–2005. <https://doi.org/10.1056/NEJMoal206809>.
- Bock, Rüdiger, Jörg Meier, László G. Nyúl, Joachim Hornegger, and Georg Michelson. 2010. "Glaucoma Risk Index: Automated Glaucoma Detection from Color Fundus Images." *Medical Image Analysis* 14 (3): 471–81. <https://doi.org/10.1016/j.media.2009.12.006>.
- Boppart, Stephen A., Mark E. Brezinski, Brett E. Bouma, Guillermo J. Tearney, and James G. Fujimoto. 1996. "Investigation of Developing Embryonic Morphology Using Optical Coherence Tomography." *Developmental Biology* 177 (1): 54–63. <https://doi.org/10.1006/dbio.1996.0144>.
- Bowns, I. R., K. Collins, S. J. Walters, and A. J. G. McDonagh. 2006. "Telemedicine in Dermatology: A Randomised Controlled Trial." *Health Technology Assessment (Winchester, England)* 10 (43): iii–iv, ix–xi, 1–39.
- Bradley, Andrew P. 1997. "The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms." *Pattern Recognition* 30 (7): 1145–59. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2).
- Brewer, Ann Chang, Dawnielle C. Endly, Jill Henley, Mahsa Amir, Blake P. Sampson, Jacqueline F. Moreau, and Robert P. Dellavalle. 2013. "Mobile Applications in Dermatology." *JAMA Dermatology* 149 (11): 1300–1304. <https://doi.org/10.1001/jamadermatol.2013.5517>.

- Burgansky-Eliash, Zvia, Gadi Wollstein, Richard A. Bilonick, Hiroshi Ishikawa, Larry Kagemann, and Joel S. Schuman. 2007. "Glaucoma Detection with the Heidelberg Retina Tomograph 3." *Ophthalmology* 114 (3): 466–71. <https://doi.org/10.1016/j.ophtha.2006.08.022>.
- Burlina, Philippe M., Neil Joshi, Michael Pekala, Katia D. Pacheco, David E. Freund, and Neil M. Bressler. 2017. "Automated Grading of Age-Related Macular Degeneration From Color Fundus Images Using Deep Convolutional Neural Networks." *JAMA Ophthalmology* 135 (11): 1170–76. <https://doi.org/10.1001/jamaophthalmol.2017.3782>.
- Burns, James A. 2012. "Optical Coherence Tomography: Imaging the Larynx." *Current Opinion in Otolaryngology & Head and Neck Surgery* 20 (6): 477. <https://doi.org/10.1097/MOO.0b013e3283582d7d>.
- Burr, J. M., G. Mowatt, R. Hernández, M. A. Siddiqui, J. Cook, T. Lourenco, C. Ramsay, et al. 2007. "The Clinical Effectiveness and Cost-Effectiveness of Screening for Open Angle Glaucoma: A Systematic Review and Economic Evaluation." *Health Technology Assessment (Winchester, England)* 11 (41): 1–190.
- Bussel, Igor I, Gadi Wollstein, and Joel S Schuman. 2014. "OCT for Glaucoma Diagnosis, Screening and Detection of Glaucoma Progression." *The British Journal of Ophthalmology* 98 (Suppl 2): ii15–19. <https://doi.org/10.1136/bjophthalmol-2013-304326>.
- Carli, Paolo, Vincenzo de Giorgi, Alessandra Chiarugi, Paolo Nardini, Martin A Weinstock, Emanuele Crocetti, Marcello Stante, and Benvenuto Giannotti. 2004. "Addition of Dermoscopy to Conventional Naked-Eye Examination in Melanoma Screening: A Randomized Study." *Journal of the American Academy of Dermatology* 50 (5): 683–89. <https://doi.org/10.1016/j.jaad.2003.09.009>.
- Carlo, Talisa E. de, Andre Romano, Nadia K. Waheed, and Jay S. Duker. 2015. "A Review of Optical Coherence Tomography Angiography (OCTA)." *International Journal of Retina and Vitreous* 1 (April): 5. <https://doi.org/10.1186/s40942-015-0005-8>.
- Chan, Christina K. W., Rita A. Gangwani, Sarah M. McGhee, JinXiao Lian, and David S. H. Wong. 2015. "Cost-Effectiveness of Screening for Intermediate Age-Related Macular Degeneration during Diabetic Retinopathy Screening." *Ophthalmology* 122 (11): 2278–85. <https://doi.org/10.1016/j.ophtha.2015.06.050>.
- Chang, S. G., Bin Yu, and M. Vetterli. 2000. "Adaptive Wavelet Thresholding for Image Denoising and Compression." *IEEE Transactions on Image Processing* 9 (9): 1532–46. <https://doi.org/10.1109/83.862633>.
- Chang, Wen-Yu, Adam Huang, Chung-Yi Yang, Chien-Hung Lee, Yin-Chun Chen, Tian-Yau Wu, and Gwo-Shing Chen. 2013. "Computer-Aided Diagnosis of Skin Lesions Using Conventional Digital Photography: A Reliability and Feasibility Study." *PLOS ONE* 8 (11): e76212. <https://doi.org/10.1371/journal.pone.0076212>.
- Chen, Jing, Andreas Stahl, Ann Hellstrom, and Lois E. Smith. 2011. "Current Update on Retinopathy of Prematurity: Screening and Treatment." *Current Opinion in Pediatrics* 23 (2): 173–78. <https://doi.org/10.1097/MOP.0b013e3283423f35>.
- Chen, X., Y. Xu, D. W. Kee Wong, T. Y. Wong, and J. Liu. 2015. "Glaucoma Detection Based on Deep Convolutional Neural Network." In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 715–18. <https://doi.org/10.1109/EMBC.2015.7318462>.

- Chen, Zhe, Mengyang Liu, Michael Minneman, Laurin Ginner, Erich Hoover, Harald Sattmann, Marco Bonesi, Wolfgang Drexler, and Rainer A. Leitgeb. 2016. "Phase-Stable Swept Source OCT Angiography in Human Skin Using an Akinetic Source." *Biomedical Optics Express* 7 (8): 3032–48. <https://doi.org/10.1364/BOE.7.003032>.
- Cheng, Jie-Zhi, Dong Ni, Yi-Hong Chou, Jing Qin, Chui-Mei Tiu, Yeun-Chung Chang, Chiun-Sheng Huang, Dinggang Shen, and Chung-Ming Chen. 2016. "Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans." *Scientific Reports* 6 (April): 24454. <https://doi.org/10.1038/srep24454>.
- Cheng, Wei, Jie Qian, Zhaoyuan Cao, Xinjian Chen, and Jianhua Mo. 2017. "Dual-Beam Angular Compounding for Speckle Reduction in Optical Coherence Tomography." In *Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XXI*, 10053:100532Z. International Society for Optics and Photonics. <https://doi.org/10.1117/12.2254270>.
- Chinn, S. R., E. A. Swanson, and J. G. Fujimoto. 1997. "Optical Coherence Tomography Using a Frequency-Tunable Optical Source." *Optics Letters* 22 (5): 340–42. <https://doi.org/10.1364/OL.22.000340>.
- Chitchian, Shahab, Michael A. Fiddy, and Nathaniel M. Fried. 2009. "Denoising during Optical Coherence Tomography of the Prostate Nerves via Wavelet Shrinkage Using Dual-Tree Complex Wavelet Transform." *Journal of Biomedical Optics* 14 (1): 014031. <https://doi.org/10.1117/1.3081543>.
- Ciulla, Thomas A., Armando G. Amador, and Bernard Zinman. 2003. "Diabetic Retinopathy and Diabetic Macular Edema: Pathophysiology, Screening, and Novel Therapies." *Diabetes Care* 26 (9): 2653–64. <https://doi.org/10.2337/diacare.26.9.2653>.
- Cobb, Michael J., Joo Ha Hwang, Melissa P. Upton, Yuchuan Chen, Brant K. Oelschlager, Douglas E. Wood, Michael B. Kimmey, and Xingde Li. 2010. "Imaging of Subsquamous Barrett's Epithelium with Ultrahigh-Resolution Optical Coherence Tomography: A Histologic Correlation Study." *Gastrointestinal Endoscopy* 71 (2): 223–30. <https://doi.org/10.1016/j.gie.2009.07.005>.
- Coleman, Anne L., and Stefano Miglior. 2008. "Risk Factors for Glaucoma Onset and Progression." *Survey of Ophthalmology* 53 (6, Supplement): S3–10. <https://doi.org/10.1016/j.survophthal.2008.08.006>.
- Coleman, Hanna R, Chi-Chao Chan, Frederick L Ferris, and Emily Y Chew. 2008. "Age-Related Macular Degeneration." *The Lancet* 372 (9652): 1835–45. [https://doi.org/10.1016/S0140-6736\(08\)61759-6](https://doi.org/10.1016/S0140-6736(08)61759-6).
- Colijn, Johanna M., Gabriëlle H. S. Buitendijk, Elena Prokofyeva, Dalila Alves, Maria L. Cachulo, Anthony P. Khawaja, Audrey Cougnard-Gregoire, et al. 2017. "Prevalence of Age-Related Macular Degeneration in Europe: The Past and the Future." *Ophthalmology* 124 (12): 1753–63. <https://doi.org/10.1016/j.ophtha.2017.05.035>.
- Commission on Chronic Illness. 1957. "Chronic Illness in the United States. Volume I: Prevention of Chronic Illness." *Journal of the American Medical Association* 165 (11): 1513–1513. <https://doi.org/10.1001/jama.1957.02980290153030>.
- Coupe, P., P. Hellier, C. Kervrann, and C. Barillot. 2009. "Nonlocal Means-Based Speckle Filtering for Ultrasound Images." *IEEE Transactions on Image Processing* 18 (10): 2221–29. <https://doi.org/10.1109/TIP.2009.2024064>.
- Cuadros, Jorge, and George Bresnick. 2009. "EyePACS: An Adaptable Telemedicine System for Diabetic Retinopathy Screening

- EyePACS: An Adaptable Telemedicine System for Diabetic Retinopathy Screening.” *Journal of Diabetes Science and Technology* 3 (3): 509–16. <https://doi.org/10.1177/193229680900300315>.
- Das, Taraprasad, Rajiv Raman, Kim Ramasamy, and Padmaja Kumari Rani. 2015. “Telemedicine in Diabetic Retinopathy: Current Status and Future Directions.” *Middle East African Journal of Ophthalmology* 22 (2): 174–78. <https://doi.org/10.4103/0974-9233.154391>.
- Davis, Jesse, and Mark Goadrich. 2006. “The Relationship Between Precision-Recall and ROC Curves.” In *Proceedings of the 23rd International Conference on Machine Learning*, 233–240. ICML ’06. New York, NY, USA: ACM. <https://doi.org/10.1145/1143844.1143874>.
- Day, Alexander C., Gianluca Baio, Gus Gazzard, Catey Bunce, Augusto Azuara-Blanco, Beatriz Munoz, David S. Friedman, and Paul J. Foster. 2012. “The Prevalence of Primary Angle Closure Glaucoma in European Derived Populations: A Systematic Review.” *British Journal of Ophthalmology*, January, bjophthalmol-2011-301189. <https://doi.org/10.1136/bjophthalmol-2011-301189>.
- De Rooij, M.j.m., F.h.j. Rampen, L.j. Schouten, and H.a.m. Neumann. 1996. “Total Skin Examination during Screening for Malignant Melanoma Does Not Increase the Detection Rate.” *British Journal of Dermatology* 135 (1): 42–45. <https://doi.org/10.1046/j.1365-2133.1996.d01-930.x>.
- Deng, J., W. Dong, R. Socher, L. J. Li, Kai Li, and Li Fei-Fei. 2009. “ImageNet: A Large-Scale Hierarchical Image Database.” In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–55. <https://doi.org/10.1109/CVPR.2009.5206848>.
- Desjardins, A. E., B. J. Vakoc, G. J. Tearney, and B. E. Bouma. 2006. “Speckle Reduction in OCT Using Massively-Parallel Detection and Frequency-Domain Ranging.” *Optics Express* 14 (11): 4736–45. <https://doi.org/10.1364/OE.14.004736>.
- Diabetic Retinopathy Clinical Research Network, David J. Browning, Adam R. Glassman, Lloyd Paul Aiello, Roy W. Beck, David M. Brown, Donald S. Fong, et al. 2007. “Relationship between Optical Coherence Tomography-Measured Central Retinal Thickness and Visual Acuity in Diabetic Macular Edema.” *Ophthalmology* 114 (3): 525–36. <https://doi.org/10.1016/j.ophtha.2006.06.052>.
- Diepgen, T.I., and V. Mahler. 2002. “The Epidemiology of Skin Cancer.” *British Journal of Dermatology* 146 (April): 1–6. <https://doi.org/10.1046/j.1365-2133.146.s61.2.x>.
- Drexler, Wolfgang, and James G. Fujimoto. 2008. *Optical Coherence Tomography: Technology and Applications*. Springer Science & Business Media.
- Drexler, Wolfgang, Mengyang Liu, Abhishek Kumar, Tschackad Kamali, Angelika Unterhuber, and Rainer A. Leitgeb. 2014. “Optical Coherence Tomography Today: Speed, Contrast, and Multimodality.” *Journal of Biomedical Optics* 19 (7): 071412. <https://doi.org/10.1117/1.JBO.19.7.071412>.
- Duan, Lian, Hee Yoon Lee, Gary Lee, Monica Agrawal, Gennifer T. Smith, and Audrey K. Ellerbee. 2014. “Single-Shot Speckle Noise Reduction by Interleaved Optical Coherence Tomography.” *Journal of Biomedical Optics* 19 (12): 120501. <https://doi.org/10.1117/1.JBO.19.12.120501>.
- Duchin, Kyla S., Baharak Asefzadeh, Vasiliki Poulaki, Douglas Rett, Paul Marescalchi, and Anthony Cavallerano. 2015. “Telereitinal Imaging for Detection of Referable Macular Degeneration.” *Optometry and Vision Science* 92 (6): 714. <https://doi.org/10.1097/OPX.0000000000000598>.

- Dutta, M. K., A. K. Mourya, A. Singh, M. Parthasarathi, R. Burget, and K. Riha. 2014. "Glaucoma Detection by Segmenting the Super Pixels from Fundus Colour Retinal Images." In *2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, 86–90. <https://doi.org/10.1109/MedCom.2014.7005981>.
- Edison, Karen E., Dana S. Ward, Jonathan A. Dyer, Whitney Lane, Louanne Chance, and Lanis L. Hicks. 2008. "Diagnosis, Diagnostic Confidence, and Management Concordance in Live-Interactive and Store-and-Forward Teledermatology Compared to In-Person Examination." *Telemedicine and E-Health* 14 (9): 889–95. <https://doi.org/10.1089/tmj.2008.0001>.
- Einarson, Thomas R., Colin Vicente, Márcio Machado, David Covert, Graham E. Trope, and Michael Iskedjian. 2006. "Screening for Glaucoma in Canada: A Systematic Review of the Literature." *Canadian Journal of Ophthalmology / Journal Canadien d'Ophtalmologie* 41 (6): 709–21. <https://doi.org/10.3129/i06-064>.
- El-Baz, Ayman, Garth M. Beache, Georgy Gimel'farb, Kenji Suzuki, Kazunori Okada, Ahmed Elnakib, Ahmed Soliman, and Behnoush Abdollahi. 2013. "Computer-Aided Diagnosis Systems for Lung Cancer: Challenges and Methodologies." *International Journal of Biomedical Imaging* 2013. <https://doi.org/10.1155/2013/942353>.
- Erdmann, Friederike, Joannie Lortet-Tieulent, Joachim Schüz, Hajo Zeeb, Rüdiger Greinert, Eckhard W. Breitbart, and Freddie Bray. 2013. "International Trends in the Incidence of Malignant Melanoma 1953–2008—Are Recent Generations at Higher or Lower Risk?" *International Journal of Cancer* 132 (2): 385–400. <https://doi.org/10.1002/ijc.27616>.
- Erl, Thomas. 2005. *Service-Oriented Architecture: Concepts, Technology, and Design*. Upper Saddle River, NJ: Prentice Hall Professional Technical Reference.
- Ervin, A. M., M. V. Boland, E. H. Myrowitz, J. Prince, B. Hawkins, D. Vollenweider, D. Ward, C. Suarez-Cuervo, and K. A. Robinson. 2012. "Screening for Glaucoma: Comparative Effectiveness." <http://europepmc.org/abstract/med/22649799>.
- Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." *Nature* 542 (7639): 115–18. <https://doi.org/10.1038/nature21056>.
- "European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part 1 Supported by the EGS Foundation." 2017. *British Journal of Ophthalmology* 101 (4): 1–72. <https://doi.org/10.1136/bjophthalmol-2016-EGSguideline.001>.
- Fang, Leyuan, David Cunefare, Chong Wang, Robyn H. Guymer, Shutao Li, and Sina Farsiu. 2017. "Automatic Segmentation of Nine Retinal Layer Boundaries in OCT Images of Non-Exudative AMD Patients Using Deep Learning and Graph Search." *Biomedical Optics Express* 8 (5): 2732–44. <https://doi.org/10.1364/BOE.8.002732>.
- Fawcett, Tom. 2006. "An Introduction to ROC Analysis." *Pattern Recognition Letters, ROC Analysis in Pattern Recognition*, 27 (8): 861–74. <https://doi.org/10.1016/j.patrec.2005.10.010>.
- Fercher, A. F., C. K. Hitzenberger, W. Drexler, G. Kamp, and H. Sattmann. 1993. "In Vivo Optical Coherence Tomography." *American Journal of Ophthalmology* 116 (1): 113–14.

- Fercher, A. F., C. K. Hitzenberger, G. Kamp, and S. Y. El-Zaiat. 1995. "Measurement of Intraocular Distances by Backscattering Spectral Interferometry." *Optics Communications* 117 (1): 43–48. [https://doi.org/10.1016/0030-4018\(95\)00119-S](https://doi.org/10.1016/0030-4018(95)00119-S).
- Fernández, Delia Cabrera, Harry M. Salinas, and Carmen A. Puliafito. 2005. "Automated Detection of Retinal Layer Structures on Optical Coherence Tomography Images." *Optics Express* 13 (25): 10200–216. <https://doi.org/10.1364/OPEX.13.010200>.
- Ferris, Frederick L. 1993. "How Effective Are Treatments for Diabetic Retinopathy?" *JAMA* 269 (10): 1290–91. <https://doi.org/10.1001/jama.1993.03500100088034>.
- Fingeret, Murray, Felipe A. Medeiros, Remo Susanna, and Robert N. Weinreb. 2005. "Five Rules to Evaluate the Optic Disc and Retinal Nerve Fiber Layer for Glaucoma." *Optometry - Journal of the American Optometric Association* 76 (11): 661–68. <https://doi.org/10.1016/j.optm.2005.08.029>.
- Finnane, Anna, Kathy Dallest, Monika Janda, and H. Peter Soyer. 2017. "Teledermatology for the Diagnosis and Management of Skin Cancer: A Systematic Review." *JAMA Dermatology* 153 (3): 319–27. <https://doi.org/10.1001/jamadermatol.2016.4361>.
- Fleming, A. D., S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp. 2006. "Automated Microaneurysm Detection Using Local Contrast Normalization and Local Vessel Detection." *IEEE Transactions on Medical Imaging* 25 (9): 1223–32. <https://doi.org/10.1109/TMI.2006.879953>.
- Fleming, Alan D., Keith A. Goatman, Sam Philip, Gordon J. Prescott, Peter F. Sharp, and John A. Olson. 2010. "Automated Grading for Diabetic Retinopathy: A Large-Scale Audit Using Arbitration by Clinical Experts." *British Journal of Ophthalmology* 94 (12): 1606–10. <https://doi.org/10.1136/bjo.2009.176784>.
- Fleming, Alan D., Sam Philip, Keith A. Goatman, John A. Olson, and Peter F. Sharp. 2006. "Automated Assessment of Diabetic Retinal Image Quality Based on Clarity and Field Definition." *Investigative Ophthalmology & Visual Science* 47 (3): 1120–25. <https://doi.org/10.1167/iovs.05-1155>.
- Fleming, Craig, Evelyn P. Whitlock, Tracy Beil, Barbara Smit, and Russell P. Harris. 2005. "Screening for Primary Open-Angle Glaucoma in the Primary Care Setting: An Update for the US Preventive Services Task Force." *The Annals of Family Medicine* 3 (2): 167–70. <https://doi.org/10.1370/afm.293>.
- Fluss, Ronen, David Faraggi, and Benjamin Reiser. 2005. "Estimation of the Youden Index and Its Associated Cutoff Point." *Biometrical Journal* 47 (4): 458–72. <https://doi.org/10.1002/bimj.200410135>.
- Forsea, AM, EM Carstea, L Ghervase, C Giurcaneanu, and G Pavelescu. 2010. "Clinical Application of Optical Coherence Tomography for the Imaging of Non-Melanocytic Cutaneous Tumors: A Pilot Multi-Modal Study." *Journal of Medicine and Life* 3 (4): 381–89.
- Foster, Paul J., David C. Broadway, David F. Garway-Heath, Jennifer L. Y. Yip, Robert Luben, Shabina Hayat, Nichola Dalzell, Nicholas J. Wareham, and Kay-Tee Khaw. 2011. "Intraocular Pressure and Corneal Biomechanics in an Adult British Population: The EPIC-Norfolk Eye Study." *Investigative Ophthalmology & Visual Science* 52 (11): 8179–85. <https://doi.org/10.1167/iovs.11-7853>.
- Freer, Timothy W., and Michael J. Ulissey. 2001. "Screening Mammography with Computer-Aided Detection: Prospective Study of 12,860 Patients in a Community Breast Center." *Radiology* 220 (3): 781–86. <https://doi.org/10.1148/radiol.2203001282>.

- Frost, S., Y. Kanagasingam, H. Sohrabi, J. Vignarajan, P. Bourgeat, O. Salvado, V. Villemagne, et al. 2013. "Retinal Vascular Biomarkers for Early Detection and Monitoring of Alzheimer's Disease." *Translational Psychiatry* 3 (2): e233. <https://doi.org/10.1038/tp.2012.150>.
- Fu, H., J. Cheng, Y. Xu, C. Zhang, D. W. K. Wong, J. Liu, and X. Cao. 2018. "Disc-Aware Ensemble Network for Glaucoma Screening from Fundus Image." *IEEE Transactions on Medical Imaging*, 1–1. <https://doi.org/10.1109/TMI.2018.2837012>.
- Fu, Huazhu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, and Jiang Liu. 2016. "DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field." In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, 132–39. Lecture Notes in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-46723-8_16.
- Fujimoto, J. G., M. E. Brezinski, G. J. Tearney, S. A. Boppart, B. Bouma, M. R. Hee, J. F. Southern, and E. A. Swanson. 1995. "Optical Biopsy and Imaging Using Optical Coherence Tomography." *Nature Medicine* 1 (9): 970–72.
- Fukushima, Kunihiro, and Sei Miyake. 1982. "Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition." In *Competition and Cooperation in Neural Nets*, 267–85. Lecture Notes in Biomathematics. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-46466-9_18.
- Fumero, F., S. Alayon, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez. 2011. "RIM-ONE: An Open Retinal Image Database for Optic Nerve Evaluation." In *2011 24th International Symposium on Computer-Based Medical Systems (CBMS)*, 1–6. <https://doi.org/10.1109/CBMS.2011.5999143>.
- Gambichler, Thilo, Hyun, Moussa, Tomi, S Boms, P Altmeyer, K Hoffmann, and A Kreuter. 2007. "Optical Coherence Tomography of Cutaneous Lupus Erythematosus Correlates with Histopathology." *Lupus* 16 (1): 35–38. <https://doi.org/10.1177/0961203306074773>.
- Gambichler, Thilo, Alexej Orlikov, Remus Vasa, Georg Moussa, Klaus Hoffmann, Markus Stücker, Peter Altmeyer, and Falk G. Bechara. 2007. "In Vivo Optical Coherence Tomography of Basal Cell Carcinoma." *Journal of Dermatological Science* 45 (3): 167–73. <https://doi.org/10.1016/j.jdermsci.2006.11.012>.
- Gambichler, Thilo, Azem Pljakic, and Lutz Schmitz. 2015. "Recent Advances in Clinical Application of Optical Coherence Tomography of Human Skin." *Clinical, Cosmetic and Investigational Dermatology* 8 (July): 345–54. <https://doi.org/10.2147/CCID.S69119>.
- García, María, María I. López, Daniel Álvarez, and Roberto Hornero. 2010. "Assessment of Four Neural Network Based Classifiers to Automatically Detect Red Lesions in Retinal Images." *Medical Engineering and Physics* 32 (10): 1085–93. <https://doi.org/10.1016/j.medengphy.2010.07.014>.
- Garg, Seema, and Richard M. Davis. 2009. "Diabetic Retinopathy Screening Update." *Clinical Diabetes* 27 (4): 140–45. <https://doi.org/10.2337/diaclin.27.4.140>.
- Goldschmidt, Leonard, and Gregory L. Goodrich. 2004. "Development and Evaluation of a Point-of-Care Interactive Patient Education Kiosk." *Journal of Telemedicine and Telecare* 10 (1_suppl): 30–32. <https://doi.org/10.1258/1357633042614159>.
- Gopinath, Karthik, Samrudhithi B. Rangrej, and Jayanthi Sivaswamy. 2018. "A Deep Learning Framework for Segmentation of Retinal Layers from OCT Images." *ArXiv:1806.08859 [Cs]*, June. <http://arxiv.org/abs/1806.08859>.

- Gossage, Kirk W., Cynthia M. Smith, Elizabeth M. Kanter, Lida P. Hariri, Alice L. Stone, Jeffrey J. Rodriguez, Stuart K. Williams, and Jennifer K. Barton. 2006. "Texture Analysis of Speckle in Optical Coherence Tomography Images of Tissue Phantoms." *Physics in Medicine & Biology* 51 (6): 1563. <https://doi.org/10.1088/0031-9155/51/6/014>.
- Götzinger, Erich, Michael Pircher, Bernhard Baumann, Tilman Schmoll, Harald Sattmann, Rainer A. Leitgeb, and Christoph K. Hitzenberger. 2011. "Speckle Noise Reduction in High Speed Polarization Sensitive Spectral Domain Optical Coherence Tomography." *Optics Express* 19 (15): 14568–84. <https://doi.org/10.1364/OE.19.014568>.
- Greaves, N.s., B. Benatar, S. Whiteside, T. Alonso-Rasgado, M. Baguneid, and A. Bayat. 2014. "Optical Coherence Tomography: A Reliable Alternative to Invasive Histological Assessment of Acute Wound Healing in Human Skin?" *British Journal of Dermatology* 170 (4): 840–50. <https://doi.org/10.1111/bjd.12786>.
- Grinsven, M. J. J. P. van, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez. 2016. "Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images." *IEEE Transactions on Medical Imaging* 35 (5): 1273–84. <https://doi.org/10.1109/TMI.2016.2526689>.
- Grinsven, Mark J. J. P. van, Yara T. E. Lechanteur, Johannes P. H. van de Ven, Bram van Ginneken, Carel B. Hoyng, Thomas Theelen, and Clara I. Sánchez. 2013. "Automatic Drusen Quantification and Risk Assessment of Age-Related Macular Degeneration on Color Fundus Images." *Investigative Ophthalmology & Visual Science* 54 (4): 3019–27. <https://doi.org/10.1167/iovs.12-11449>.
- Gulshan, Varun, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, et al. 2016. "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs." *JAMA* 316 (22): 2402–10. <https://doi.org/10.1001/jama.2016.17216>.
- Gupta, S.C., Subodh Kumar Sinha, and Abhishek B. Dagar. 2013. "Evaluation of the Effectiveness of Diagnostic & Management Decision by Teleophthalmology Using Indigenous Equipment in Comparison with In-Clinic Assessment of Patients." *The Indian Journal of Medical Research* 138 (4): 531–35.
- Haines, L., A. R. Fielder, H. Baker, and A. R. Wilkinson. 2005. "UK Population Based Study of Severe Retinopathy of Prematurity: Screening, Treatment, and Outcome." *Archives of Disease in Childhood - Fetal and Neonatal Edition* 90 (3): F240–FF244. <https://doi.org/10.1136/adc.2004.057570>.
- Haleem, Muhammad Salman, Liangxiu Han, Jano van Hemert, and Baihua Li. 2013. "Automatic Extraction of Retinal Features from Colour Retinal Images for Glaucoma Diagnosis: A Review." *Computerized Medical Imaging and Graphics* 37 (7): 581–96. <https://doi.org/10.1016/j.compmedimag.2013.09.005>.
- Hark, Lisa A., L. Jay Katz, Jonathan S. Myers, Michael Waisbourd, Deiana Johnson, Laura T. Pizzi, Benjamin E. Leiby, et al. 2017. "Philadelphia Telemedicine Glaucoma Detection and Follow-up Study: Methods and Screening Results." *American Journal of Ophthalmology* 181 (September): 114–24. <https://doi.org/10.1016/j.ajo.2017.06.024>.
- Hatta, Waku, Kaname Uno, Tomoyuki Koike, Satoshi Yokosawa, Katsunori Iijima, Akira Imatani, and Tooru Shimosegawa. 2010. "Optical Coherence Tomography for the

- Staging of Tumor Infiltration in Superficial Esophageal Squamous Cell Carcinoma.” *Gastrointestinal Endoscopy* 71 (6): 899–906. <https://doi.org/10.1016/j.gie.2009.11.052>.
- He, Mingguang, Wenyong Huang, Yingfeng Zheng, Li Huang, and Leon B. Ellwein. 2007. “Refractive Error and Visual Impairment in School Children in Rural Southern China.” *Ophthalmology* 114 (2): 374–382.e1. <https://doi.org/10.1016/j.ophtha.2006.08.020>.
- Healey, Paul R., Anne J. Lee, Tin Aung, Tien Y. Wong, and Paul Mitchell. 2010. “Diagnostic Accuracy of the Heidelberg Retina Tomograph for Glaucoma: A Population-Based Assessment.” *Ophthalmology* 117 (9): 1667–73. <https://doi.org/10.1016/j.ophtha.2010.07.001>.
- Health Departments of the United Kingdom. 2000. “Second Report of the UK National Screening Committee.” Health Departments of the United Kingdom. <http://aogm.org.mo/assets/Uploads/aogm/Guidelines/RCOG---UK/Second-report-of-UK-NSC.pdf>.
- Hee, Michael R., Carmen A. Puliafito, Carlton Wong, Jay S. Duker, Elias Reichel, Joel S. Schuman, Eric A. Swanson, and James G. Fujimoto. 1995. “Optical Coherence Tomography of Macular Holes.” *Ophthalmology* 102 (5): 748–56. [https://doi.org/10.1016/S0161-6420\(95\)30959-1](https://doi.org/10.1016/S0161-6420(95)30959-1).
- Heshka, Jodi T., Crystal Palleschi, Heather Howley, Brenda Wilson, and Philip S. Wells. 2008. “A Systematic Review of Perceived Risks, Psychological and Behavioral Impacts of Genetic Testing.” *Genetics in Medicine* 10 (1): 19–32. <https://doi.org/10.1097/GIM.0b013e31815f524f>.
- Hitzl, W., C. Ortner, K. Hornykewycz, G. Grabner, and H. A. Reitsamer. 2006. “Resource Use and Costs for a Glaucoma Screening Program in Austria: An 8-Year Review.: A Cost-Consequence Analysis Based on the Salzburg-Moorfields Collaborative Glaucoma Study.” *European Journal of Ophthalmology* 16 (1): 92–99. <https://doi.org/10.1177/112067210601600115>.
- Hoehn, René, Alireza Mirshahi, Esther M. Hoffmann, Ulrike B. Kottler, Philipp S. Wild, Dagmar Laubert-Reh, and Norbert Pfeiffer. 2013. “Distribution of Intraocular Pressure and Its Association with Ocular Features and Cardiovascular Risk Factors: The Gutenberg Health Study.” *Ophthalmology* 120 (5): 961–68. <https://doi.org/10.1016/j.ophtha.2012.10.031>.
- Holland, Stewart, and Masseria. 2006. “Policy Brief: Screening in Europe.” Copenhagen : WHO Regional Office for Europe. <http://apps.who.int/iris/bitstream/10665/107742/1/E88698.pdf>.
- Holland, Walter Werner, and Susie Stewart. 2005. *Screening in Disease Prevention: What Works?* Radcliffe Publishing.
- Huang, David, Eric A. Swanson, Charles P. Lin, Joel S. Schuman, William G. Stinson, Warren Chang, Michael R. Hee, et al. 1991. “Optical Coherence Tomography.” *Science (New York, N.Y.)* 254 (5035): 1178–81.
- Huang, T., G. Yang, and G. Tang. 1979. “A Fast Two-Dimensional Median Filtering Algorithm.” *IEEE Transactions on Acoustics, Speech, and Signal Processing* 27 (1): 13–18. <https://doi.org/10.1109/TASSP.1979.1163188>.
- Hunter, NorwoodCaputo. 1994. “The Portable Dynamic Fundus Instrument: Uses in Telemedicine and Research.” In . <https://ntrs.nasa.gov/search.jsp?R=19940029141>.

- Hussain, Alia Arif, Lotte Themstrup, and Gregor Borut Ernst Jemec. 2015. "Optical Coherence Tomography in the Diagnosis of Basal Cell Carcinoma." *Archives of Dermatological Research* 307 (1): 1–10. <https://doi.org/10.1007/s00403-014-1498-y>.
- Huynh, Benjamin Q., Hui Li, and Maryellen L. Giger. 2016. "Digital Mammographic Tumor Classification Using Transfer Learning from Deep Convolutional Neural Networks." *Journal of Medical Imaging* 3 (3). <https://doi.org/10.1117/1.JMI.3.3.034501>.
- Ian Levitt. 2003. "TB, Glasgow and the Mass Radiography Campaign in the Nineteen Fifties: Preserving a Democratic Health Service." 2003. https://www.gla.ac.uk/media/media_148456_en.pdf.
- Iftimia, Nicusor, Brett E. Bouma, and Guillermo J. Tearney. 2003. "Speckle Reduction in Optical Coherence Tomography By." *Journal of Biomedical Optics* 8 (2): 260–64. <https://doi.org/10.1117/1.1559060>.
- International Agency for Research on Cancer, CPO Piemonte and University Hospital "Città della Salute e della Scienza," and Mass Screening Registry/ Finnish Cancer Registry. 2017. "Cancer Screening in the European Union (2017)." European Union Public Health Programme. https://ec.europa.eu/health/sites/health/files/major_chronic_diseases/docs/2017_cancerscreening_2ndreportimplementation_en.pdf.
- Iyatomi, Hitoshi, Hiroshi Oka, M. Emre Celebi, Masahiro Hashimoto, Masafumi Hagiwara, Masaru Tanaka, and Koichi Ogawa. 2008. "An Improved Internet-Based Melanoma Screening System with Dermatologist-like Tumor Area Extraction Algorithm." *Computerized Medical Imaging and Graphics* 32 (7): 566–79. <https://doi.org/10.1016/j.compmedimag.2008.06.005>.
- Jaffe, Glenn J., and Joseph Caprioli. 2004. "Optical Coherence Tomography to Detect and Manage Retinal Disease and Glaucoma." *American Journal of Ophthalmology* 137 (1): 156–69. [https://doi.org/10.1016/S0002-9394\(03\)00792-X](https://doi.org/10.1016/S0002-9394(03)00792-X).
- Jennifer K. Barton, Kirk W. Gossage, Wei Xu, James R. Ranger-Moore, Kathylynn Saboda, Christine A. Brooks, Laura D. Duckett, Stuart J. Salasche, James A. Warneke, and David S. Alberts. 2003. "Investigating Sun-Damaged Skin and Actinic Keratosis with Optical Coherence Tomography: A Pilot Study." *Technology in Cancer Research & Treatment* 2 (6): 525–35. <https://doi.org/10.1177/153303460300200605>.
- Jia, Yangqing, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. 2014. "Caffe: Convolutional Architecture for Fast Feature Embedding." In *Proceedings of the 22Nd ACM International Conference on Multimedia*, 675–678. MM '14. New York, NY, USA: ACM. <https://doi.org/10.1145/2647868.2654889>.
- Jonas, Jost B., Wido M. Budde, and Songhomitra Panda-Jonas. 1999. "Ophthalmoscopic Evaluation of the Optic Nerve Head." *Survey of Ophthalmology* 43 (4): 293–320. [https://doi.org/10.1016/S0039-6257\(98\)00049-6](https://doi.org/10.1016/S0039-6257(98)00049-6).
- Jong, Antina de, Wybo J. Dondorp, Christine E. M. de Die-Smulders, Suzanne G. M. Frints, and Guido M. W. R. de Wert. 2010. "Non-Invasive Prenatal Testing: Ethical Issues Explored." *European Journal of Human Genetics* 18 (3): 272–77. <https://doi.org/10.1038/ejhg.2009.203>.
- Joshi, G. D., J. Sivaswamy, and S. R. Krishnadas. 2011. "Optic Disk and Cup Segmentation From Monocular Color Retinal Images for Glaucoma Assessment." *IEEE*

- Transactions on Medical Imaging* 30 (6): 1192–1205. <https://doi.org/10.1109/TMI.2011.2106509>.
- Jung, W., J. Kim, M. Jeon, E. J. Chaney, C. N. Stewart, and S. A. Boppart. 2011. “Handheld Optical Coherence Tomography Scanner for Primary Care Diagnostics.” *IEEE Transactions on Biomedical Engineering* 58 (3): 741–44. <https://doi.org/10.1109/TBME.2010.2096816>.
- Kamdeu Fansi, Alvine A., Gisle Li, and Paul J. Harasymowycz. 2011. “The Validity of Screening For Open-Angle Glaucoma in High-Risk Populations With Single-Test Screening Mode Frequency Doubling Technology Perimetry (FDT).” *Journal of Glaucoma* 20 (3): 167. <https://doi.org/10.1097/IJG.0b013e3181d9cb7c>.
- Karri, S. P. K., Debjani Chakraborty, and Jyotirmoy Chatterjee. 2017. “Transfer Learning Based Classification of Optical Coherence Tomography Images with Diabetic Macular Edema and Dry Age-Related Macular Degeneration.” *Biomedical Optics Express* 8 (2): 579–92. <https://doi.org/10.1364/BOE.8.000579>.
- Kassam, Faazil, Kanagasingam Yogesana, Enitan Sogbesan, Louis R. Pasquale, and Karim F. Damji. 2013. “Teleglaucoma: Improving Access and Efficiency for Glaucoma Care.” *Middle East African Journal of Ophthalmology* 20 (2): 142–49. <https://doi.org/10.4103/0974-9233.110619>.
- Kassianos, A.p., J.d. Emery, P. Murchie, and F.m. Walter. 2015. “Smartphone Applications for Melanoma Detection by Community, Patient and Generalist Clinician Users: A Review.” *British Journal of Dermatology* 172 (6): 1507–18. <https://doi.org/10.1111/bjd.13665>.
- Khan, F., S. A. Khan, U. U. Yasin, I. ul Haq, and U. Qamar. 2013. “Detection of Glaucoma Using Retinal Fundus Images.” In *The 6th 2013 Biomedical Engineering International Conference*, 1–5. <https://doi.org/10.1109/BMEiCon.2013.6687674>.
- Kiage, Dan, Irfan N. Kherani, Stephen Gichuhi, Karim F. Damji, and Muindi Nyenze. 2013. “The Muranga Teleophthalmology Study: Comparison of Virtual (Teleglaucoma) with in-Person Clinical Assessment to Diagnose Glaucoma.” *Middle East African Journal of Ophthalmology* 20 (2): 150–57. <https://doi.org/10.4103/0974-9233.110604>.
- Kirkpatrick, Sean J., Ruikang K. Wang, and Donald D. Duncan. 2006. “OCT-Based Elastography for Large and Small Deformations.” *Optics Express* 14 (24): 11585–97. <https://doi.org/10.1364/OE.14.011585>.
- Klein, Ronald, Michael D. Knudtson, Kristine E. Lee, Ronald Gangnon, and Barbara E. K. Klein. 2008. “The Wisconsin Epidemiologic Study of Diabetic Retinopathy XXII: The Twenty-Five-Year Progression of Retinopathy in Persons with Type 1 Diabetes.” *Ophthalmology* 115 (11): 1859–68. <https://doi.org/10.1016/j.ophtha.2008.08.023>.
- Kooi, Thijs, Geert Litjens, Bram van Ginneken, Albert Gubern-Mrida, Clara I. Snchez, Ritse Mann, Ard den Heeten, and Nico Karssemeijer. 2017. “Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions.” *Medical Image Analysis* 35 (January): 303–12. <https://doi.org/10.1016/j.media.2016.07.007>.
- Krishnan, M Muthu Rama, and Oliver Faust. 2013. “AUTOMATED GLAUCOMA DETECTION USING HYBRID FEATURE EXTRACTION IN RETINAL FUNDUS IMAGES.” *Journal of Mechanics in Medicine and Biology* 13 (01): 1350011. <https://doi.org/10.1142/S0219519413500115>.
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton. 2012. “ImageNet Classification with Deep Convolutional Neural Networks.” In *Advances in Neural Information*

- Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, 1097–1105. Curran Associates, Inc. <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- Kumar, S., M.-L. Tay-Kearney, I. J. Constable, and K. Yogesan. 2005. “Internet Based Ophthalmology Service: Impact Assessment.” *British Journal of Ophthalmology* 89 (10): 1382–83. <https://doi.org/10.1136/bjo.2005.072579>.
- Kumar, S., K. Yogesan, B. Hudson, M.-L. Tay-Kearney, and I. J. Constable. 2006. “Emergency Eye Care in Rural Australia: Role of Internet.” *Eye* 20 (12): 1342–44. <https://doi.org/10.1038/sj.eye.6702104>.
- Kvarnström, Gun, Peter Jakobsson, and Gunnar Lennerstrand. 2001. “Visual Screening of Swedish Children: An Ophthalmological Evaluation.” *Acta Ophthalmologica Scandinavica* 79 (3): 240–44. <https://doi.org/10.1034/j.1600-0420.2001.790306.x>.
- Lamel, Sonia A., Kristin M. Haldeman, Haines Ely, Carrie L. Kovarik, Hon Pak, and April W. Armstrong. 2012. “Application of Mobile Tele dermatology for Skin Cancer Screening.” *Journal of the American Academy of Dermatology* 67 (4): 576–81. <https://doi.org/10.1016/j.jaad.2011.11.957>.
- Larsen, Michael, Jannik Godt, Nicolai Larsen, Henrik Lund-Andersen, Anne Katrin Sjølie, Elisabet Agardh, Helle Kalm, Michael Grunkin, and David R. Owens. 2003. “Automated Detection of Fundus Photographic Red Lesions in Diabetic Retinopathy.” *Investigative Ophthalmology & Visual Science* 44 (2): 761–66. <https://doi.org/10.1167/iovs.02-0418>.
- Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. “Gradient-Based Learning Applied to Document Recognition.” *Proceedings of the IEEE* 86 (11): 2278–2324. <https://doi.org/10.1109/5.726791>.
- Lee, J. S., J. H. Wen, T. L. Ainsworth, K. S. Chen, and A. J. Chen. 2009. “Improved Sigma Filter for Speckle Filtering of SAR Imagery.” *IEEE Transactions on Geoscience and Remote Sensing* 47 (1): 202–13. <https://doi.org/10.1109/TGRS.2008.2002881>.
- Lee, Jong-Sen. 1983. “Digital Image Smoothing and the Sigma Filter.” *Computer Vision, Graphics, and Image Processing* 24 (2): 255–69. [https://doi.org/10.1016/0734-189X\(83\)90047-6](https://doi.org/10.1016/0734-189X(83)90047-6).
- Lee, Peng, Wanrong Gao, and Xianling Zhang. 2011. “Speckle Properties of the Logarithmically Transformed Signal in Optical Coherence Tomography.” *JOSA A* 28 (4): 517–22. <https://doi.org/10.1364/JOSAA.28.000517>.
- Leff, A., and J. T. Rayfield. 2001. “Web-Application Development Using the Model/View/Controller Design Pattern.” In *Proceedings Fifth IEEE International Enterprise Distributed Object Computing Conference*, 118–27. <https://doi.org/10.1109/EDOC.2001.950428>.
- Leung, Christopher Kai-shun, Carol Yim Lui Cheung, Robert N. Weinreb, Kunliang Qiu, Shu Liu, Haitao Li, Guihua Xu, et al. 2010. “Evaluation of Retinal Nerve Fiber Layer Progression in Glaucoma: A Study on Optical Coherence Tomography Guided Progression Analysis.” *Investigative Ophthalmology & Visual Science* 51 (1): 217–22. <https://doi.org/10.1167/iovs.09-3468>.
- Li, Bo, Anne-Marie Powell, Philip L. Hooper, and Thomas G. Sheidow. 2015. “Prospective Evaluation of Teleophthalmology in Screening and Recurrence Monitoring of Neovascular Age-Related Macular Degeneration: A Randomized Clinical Trial.” *JAMA Ophthalmology* 133 (3): 276–82. <https://doi.org/10.1001/jamaophthalmol.2014.5014>.

- Li, Gisèle, Alvine Kamdeu Fansi, Jean-François Boivin, Lawrence Joseph, and Paul Harasymowycz. 2010. "Screening for Glaucoma in High-Risk Populations Using Optical Coherence Tomography." *Ophthalmology* 117 (3): 453–61. <https://doi.org/10.1016/j.ophtha.2009.07.033>.
- Li, Helen K., Mark Horton, Sven-Erik Bursell, Jerry Cavallerano, Ingrid Zimmer-Galler, Mathew Tennant, Michael Abramoff, et al. 2011. "Telehealth Practice Recommendations for Diabetic Retinopathy, Second Edition." *Telemedicine and E-Health* 17 (10): 814–37. <https://doi.org/10.1089/tmj.2011.0075>.
- Li, X. D., S. A. Boppart, J. Van Dam, H. Mashimo, M. Mutinga, W. Drexler, M. Klein, et al. 2000. "Optical Coherence Tomography: Advanced Technology for the Endoscopic Imaging of Barrett's Esophagus." *Endoscopy* 32 (12): 921–30. <https://doi.org/10.1055/s-2000-9626>.
- Liang, Z., D. W. K. Wong, J. Liu, K. L. Chan, and T. Y. Wong. 2010. "Towards Automatic Detection of Age-Related Macular Degeneration in Retinal Fundus Images." In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 4100–4103. <https://doi.org/10.1109/IEMBS.2010.5627289>.
- Liesenfeld, B., E. Kohner, W. Piehlmeier, S. Kluthe, S. Aldington, M. Porta, T. Bek, et al. 2000. "A Telemedical Approach to the Screening of Diabetic Retinopathy: Digital Fundus Photography." *Diabetes Care* 23 (3): 345–48. <https://doi.org/10.2337/diacare.23.3.345>.
- Lim, Teik-Cheng, Subhagata Chattopadhyay, and U. Rajendra Acharya. 2012. "A Survey and Comparative Study on the Instruments for Glaucoma Detection." *Medical Engineering & Physics* 34 (2): 129–39. <https://doi.org/10.1016/j.medengphy.2011.07.030>.
- Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. 2017. "A Survey on Deep Learning in Medical Image Analysis." *Medical Image Analysis* 42 (December): 60–88. <https://doi.org/10.1016/j.media.2017.07.005>.
- Liu, S., S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, M. J. Fulham, and ADNI. 2015. "Multimodal Neuroimaging Feature Learning for Multiclass Diagnosis of Alzheimer #x0027;s Disease." *IEEE Transactions on Biomedical Engineering* 62 (4): 1132–40. <https://doi.org/10.1109/TBME.2014.2372011>.
- Lodwick, Gwilym S., Cosmo L. Haun, Walton E. Smith, Roy F. Keller, and Eddie D. Robertson. 1963. "Computer Diagnosis of Primary Bone Tumors." *Radiology* 80 (2): 273–75. <https://doi.org/10.1148/80.2.273>.
- Lorenz, Birgit, Katerina Spasovska, Heike Elflein, and Nico Schneider. 2009. "Wide-Field Digital Imaging Based Telemedicine for Screening for Acute Retinopathy of Prematurity (ROP). Six-Year Results of a Multicentre Field Study." *Graefe's Archive for Clinical and Experimental Ophthalmology = Albrecht Von Graefes Archiv Fur Klinische Und Experimentelle Ophthalmologie* 247 (9): 1251–62. <https://doi.org/10.1007/s00417-009-1077-7>.
- Lou, L., E. A. Sickles, H. K. Huang, D. Hoogstrate, Fei Cao, Jun Wang, and M. Jahangiri. 1997. "Full-Field Direct Digital Telemammography: Technical Components, Study Protocols, and Preliminary Results." *IEEE Transactions on Information Technology in Biomedicine* 1 (4): 270–78. <https://doi.org/10.1109/4233.681171>.

- Lowitt, Mark H., Irving I. Kessler, C. Lisa Kauffman, Frank J. Hooper, Eliot Siegel, and Joseph W. Burnett. 1998. "Teledermatology and In-Person Examinations: A Comparison of Patient and Physician Perceptions and Diagnostic Agreement." *Archives of Dermatology* 134 (4): 471–76. <https://doi.org/10.1001/archderm.134.4.471>.
- Mahapatra, Dwarikanath, Pallab K. Roy, Suman Sedai, and Rahil Garnavi. 2016. "Retinal Image Quality Classification Using Saliency Maps and CNNs." In *Machine Learning in Medical Imaging*, 172–79. Lecture Notes in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-47157-0_21.
- Mahendran, R., M. J. D. Goodfield, and R. A. Sheehan-Dare. 2005. "An Evaluation of the Role of a Store-and-Forward Teledermatology System in Skin Cancer Diagnosis and Management." *Clinical and Experimental Dermatology* 30 (3): 209–14. <https://doi.org/10.1111/j.1365-2230.2005.01735.x>.
- Maheshwari, S., R. B. Pachori, and U. R. Acharya. 2017. "Automated Diagnosis of Glaucoma Using Empirical Wavelet Transform and Correntropy Features Extracted From Fundus Images." *IEEE Journal of Biomedical and Health Informatics* 21 (3): 803–13. <https://doi.org/10.1109/JBHI.2016.2544961>.
- Maier, David. 1983. *Theory of Relational Databases*. Computer Science Pr.
- Mane, V. M., R. B. Kawadiwale, and D. V. Jadhav. 2015. "Detection of Red Lesions in Diabetic Retinopathy Affected Fundus Images." In *2015 IEEE International Advance Computing Conference (IACC)*, 56–60. <https://doi.org/10.1109/IADCC.2015.7154668>.
- Manjunath, Varsha, Jordana Goren, James G. Fujimoto, and Jay S. Duker. 2011. "Analysis of Choroidal Thickness in Age-Related Macular Degeneration Using Spectral-Domain Optical Coherence Tomography." *American Journal of Ophthalmology* 152 (4): 663–68. <https://doi.org/10.1016/j.ajo.2011.03.008>.
- Mariampillai, Adrian, Beau A. Standish, Eduardo H. Moriyama, Mamta Khurana, Nigel R. Munce, Michael K. K. Leung, James Jiang, et al. 2008. "Speckle Variance Detection of Microvasculature Using Swept-Source Optical Coherence Tomography." *Optics Letters* 33 (13): 1530–32. <https://doi.org/10.1364/OL.33.001530>.
- Marks, Daniel L., Tyler S. Ralston, and Stephen A. Boppart. 2005. "Speckle Reduction by I-Divergence Regularization in Optical Coherence Tomography." *JOSA A* 22 (11): 2366–71. <https://doi.org/10.1364/JOSAA.22.002366>.
- Marmot, M. G., D. G. Altman, D. A. Cameron, J. A. Dewar, S. G. Thompson, and M. Wilcox. 2013. "The Benefits and Harms of Breast Cancer Screening: An Independent Review." *British Journal of Cancer* 108 (11): 2205–40. <https://doi.org/10.1038/bjc.2013.177>.
- Marrugo, Andrés G., and María S. Millán. 2011. "Retinal Image Analysis: Preprocessing and Feature Extraction." *Journal of Physics: Conference Series* 274 (1): 012039. <https://doi.org/10.1088/1742-6596/274/1/012039>.
- Masood, Ammara, and Adel Ali Al-Jumaily. 2013. "Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques and Algorithms." *International Journal of Biomedical Imaging*, January, 1–22. <https://doi.org/10.1155/2013/323268>.
- Maul, Eugenio A., and Henry D. Jampel. 2010. "Glaucoma Screening in the Real World." *Ophthalmology* 117 (9): 1665–66. <https://doi.org/10.1016/j.ophtha.2009.11.001>.
- Mayer, Markus A., Anja Borsdorf, Martin Wagner, Joachim Hornegger, Christian Y. Mardin, and Ralf P. Tornow. 2012. "Wavelet Denoising of Multiframe Optical Coherence

- Tomography Data.” *Biomedical Optics Express* 3 (3): 572–89. <https://doi.org/10.1364/BOE.3.000572>.
- Medeiros, Felipe A., Linda M. Zangwill, Christopher Bowd, Roberto M. Vessani, Remo Susanna, and Robert N. Weinreb. 2005. “Evaluation of Retinal Nerve Fiber Layer, Optic Nerve Head, and Macular Thickness Measurements for Glaucoma Detection Using Optical Coherence Tomography.” *American Journal of Ophthalmology* 139 (1): 44–55. <https://doi.org/10.1016/j.ajo.2004.08.069>.
- Medeiros, Felipe A., Linda M. Zangwill, Christopher Bowd, and Robert N. Weinreb. 2004. “Comparison of the GDx VCC Scanning Laser Polarimeter, HRT II Confocal Scanning Laser Ophthalmoscope, and Stratus OCT Optical Coherence Tomograph for the Detection of Glaucoma.” *Archives of Ophthalmology* 122 (6): 827–37. <https://doi.org/10.1001/archophth.122.6.827>.
- Melendez, Jaime, Clara I. Sánchez, Rick H. H. M. Philipsen, Pragnya Maduskar, Rodney Dawson, Grant Theron, Keertan Dheda, and Bram van Ginneken. 2016. “An Automated Tuberculosis Screening Strategy Combining X-Ray-Based Computer-Aided Detection and Clinical Information.” *Scientific Reports* 6 (April): 25265. <https://doi.org/10.1038/srep25265>.
- Men, J., Y. Huang, J. Solanki, X. Zeng, A. Alex, J. Jerwick, Z. Zhang, R. E. Tanzi, A. Li, and C. Zhou. 2016. “Optical Coherence Tomography for Brain Imaging and Developmental Biology.” *IEEE Journal of Selected Topics in Quantum Electronics* 22 (4): 120–32. <https://doi.org/10.1109/JSTQE.2015.2513667>.
- Meyers, Phillip H., Charles M. Nice, Hal C. Becker, Wilson J. Nettleton, James W. Sweeney, and George R. Meckstroth. 1964. “Automated Computer Analysis of Radiographic Images.” *Radiology* 83 (6): 1029–34. <https://doi.org/10.1148/83.6.1029>.
- Michel, Ross G., Gary T. Kinasewitz, Kar-Ming Fung, and Jean I. Keddissi. 2010. “Optical Coherence Tomography as an Adjunct to Flexible Bronchoscopy in the Diagnosis of Lung Cancer: A Pilot Study.” *CHEST* 138 (4): 984–88. <https://doi.org/10.1378/chest.10-0753>.
- Mitchell, Paul, Wayne Smith, Karin Attebo, and Paul R. Healey. 1996. “Prevalence of Open-Angle Glaucoma in Australia: The Blue Mountains Eye Study.” *Ophthalmology* 103 (10): 1661–69. [https://doi.org/10.1016/S0161-6420\(96\)30449-1](https://doi.org/10.1016/S0161-6420(96)30449-1).
- Mobadersany, Pooya, Safoora Yousefi, Mohamed Amgad, David A. Gutman, Jill S. Barnholtz-Sloan, José E. Velázquez Vega, Daniel J. Brat, and Lee A. D. Cooper. 2018. “Predicting Cancer Outcomes from Histology and Genomics Using Convolutional Networks.” *Proceedings of the National Academy of Sciences*, March, 201717139. <https://doi.org/10.1073/pnas.1717139115>.
- Mogensen, M., Ha Morsy, Bm Nurnberg, and Gbe Jemec. 2008. “Optical Coherence Tomography Imaging of Bullous Diseases.” *Journal of the European Academy of Dermatology and Venereology* 22 (12): 1458–64. <https://doi.org/10.1111/j.1468-3083.2008.02917.x>.
- Mogensen, M., B.m. Nürnberg, J.l. Forman, J.b. Thomsen, L. Thrane, and G.b.e. Jemec. 2009. “In Vivo Thickness Measurement of Basal Cell Carcinoma and Actinic Keratosis with Optical Coherence Tomography and 20-MHz Ultrasound.” *British Journal of Dermatology* 160 (5): 1026–33. <https://doi.org/10.1111/j.1365-2133.2008.09003.x>.

- Mohamed, Quresh, Mark C. Gillies, and Tien Y. Wong. 2007. "Management of Diabetic Retinopathy: A Systematic Review." *JAMA* 298 (8): 902–16. <https://doi.org/10.1001/jama.298.8.902>.
- Mookiah, Muthu Rama Krishnan, U. Rajendra Acharya, Chua Kuang Chua, Choo Min Lim, E. Y. K. Ng, and Augustinus Laude. 2013. "Computer-Aided Diagnosis of Diabetic Retinopathy: A Review." *Computers in Biology and Medicine* 43 (12): 2136–55. <https://doi.org/10.1016/j.combiomed.2013.10.007>.
- Mookiah, Muthu Rama Krishnan, U. Rajendra Acharya, Chua Kuang Chua, Lim Choo Min, EYK Ng, Milind M Mushrif, and Augustinus Laude. 2013. "Automated Detection of Optic Disk in Retinal Fundus Images Using Intuitionistic Fuzzy Histon Segmentation
, Fundus Images Using Intuitionistic Fuzzy Histon Segmentation." *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 227 (1): 37–49. <https://doi.org/10.1177/0954411912458740>.
- Mookiah, Muthu Rama Krishnan, U. Rajendra Acharya, Joel E. W. Koh, Vinod Chandran, Chua Kuang Chua, Jen Hong Tan, Choo Min Lim, et al. 2014. "Automated Diagnosis of Age-Related Macular Degeneration Using Greyscale Features from Digital Fundus Images." *Computers in Biology and Medicine* 53 (October): 55–64. <https://doi.org/10.1016/j.combiomed.2014.07.015>.
- Mookiah, Muthu Rama Krishnan, U. Rajendra Acharya, Choo Min Lim, Andrea Petznick, and Jasjit S. Suri. 2012. "Data Mining Technique for Automated Diagnosis of Glaucoma Using Higher Order Spectra and Wavelet Energy Features." *Knowledge-Based Systems* 33 (September): 73–82. <https://doi.org/10.1016/j.knosys.2012.02.010>.
- Morabia, A, and F Zhang. 2004. "History of Medical Screening: From Concepts to Action." *Postgraduate Medical Journal* 80 (946): 463–69. <https://doi.org/10.1136/pgmj.2003.018226>.
- Morsy, Hanan, Soren Kamp, Lars Thrane, Nille Behrendt, Birgit Saunder, Hisham Zayan, Ensaf Abd Elmagid, and Gregor B. E. Jemec. 2010. "Optical Coherence Tomography Imaging of Psoriasis Vulgaris: Correlation with Histology and Disease Severity." *Archives of Dermatological Research* 302 (2): 105–11. <https://doi.org/10.1007/s00403-009-1000-4>.
- Mowatt, Graham, Jennifer M. Burr, Jonathan A. Cook, M. A. Rehman Siddiqui, Craig Ramsay, Cynthia Fraser, Augusto Azuara-Blanco, and Jonathan J. Deeks. 2008. "Screening Tests for Detecting Open-Angle Glaucoma: Systematic Review and Meta-Analysis." *Investigative Ophthalmology & Visual Science* 49 (12): 5373–85. <https://doi.org/10.1167/iovs.07-1501>.
- Muhammad, Hassan, Thomas J. Fuchs, Nicole De Cuir, Carlos G. De Moraes, Dana M. Blumberg, Jeffrey M. Liebmann, Robert Ritch, and Donald C. Hood. 2017. "Hybrid Deep Learning on Single Wide-Field Optical Coherence Tomography Scans Accurately Classifies Glaucoma Suspects." Text. December 2017. <https://doi.org/info:doi/10.1097/IJG.0000000000000765>.
- Muramatsu, Chisako, Toshiaki Nakagawa, Akira Sawada, Yuji Hatanaka, Takeshi Hara, Tetsuya Yamamoto, and Hiroshi Fujita. 2009. "Determination of Cup-to-Disc Ratio of Optical Nerve Head for Diagnosis of Glaucoma on Stereo Retinal Fundus Image Pairs." In *Medical Imaging 2009: Computer-Aided Diagnosis*, 7260:72603L. International Society for Optics and Photonics. <https://doi.org/10.1117/12.811461>.

- Murphy, K., B. van Ginneken, A. M. R. Schilham, B. J. de Hoop, H. A. Gietema, and M. Prokop. 2009. "A Large-Scale Evaluation of Automatic Pulmonary Nodule Detection in Chest CT Using Local Image Features and k-Nearest-Neighbour Classification." *Medical Image Analysis* 13 (5): 757–70. <https://doi.org/10.1016/j.media.2009.07.001>.
- Muyoyeta, Monde, Pragnya Maduskar, Maureen Moyo, Nkatya Kasese, Deborah Milimo, Rosanna Spooner, Nathan Kapata, Laurens Hogeweg, Bram van Ginneken, and Helen Ayles. 2014. "The Sensitivity and Specificity of Using a Computer Aided Diagnosis Program for Automatically Scoring Chest X-Rays of Presumptive TB Patients Compared with Xpert MTB/RIF in Lusaka Zambia." *PLOS ONE* 9 (4): e93757. <https://doi.org/10.1371/journal.pone.0093757>.
- Narayanaswamy, Arun, Lisandro M. Sakata, Ming-Guang He, David S. Friedman, Yiong-Huak Chan, Raghavan Lavanya, Mani Baskaran, Paul J. Foster, and Tin Aung. 2010. "Diagnostic Performance of Anterior Chamber Angle Measurements for Detecting Eyes With Narrow Angles: An Anterior Segment OCT Study." *Archives of Ophthalmology* 128 (10): 1321–27. <https://doi.org/10.1001/archophthalmol.2010.231>.
- Nayak, Jagadish, Rajendra Acharya U, P. Subbanna Bhat, Nakul Shetty, and Teik-Cheng Lim. 2009. "Automated Diagnosis of Glaucoma Using Digital Fundus Images." *Journal of Medical Systems* 33 (5): 337. <https://doi.org/10.1007/s10916-008-9195-z>.
- Neubauer, Aljoscha S., Ulrich C. Welge-Lüssen, Martin J. Thiel, Claudia Alge, Siegfried G. Priglinger, Christoph Hirneiß, Michael W. Ulbig, and Anselm Kampik. 2003. "Tele-Screening for Diabetic Retinopathy With the Retinal Thickness Analyzer." *Diabetes Care* 26 (10): 2890–97. <https://doi.org/10.2337/diacare.26.10.2890>.
- Nickells, Robert W., Gareth R. Howell, Ileana Soto, and Simon W.M. John. 2012. "Under Pressure: Cellular and Molecular Responses During Glaucoma, a Common Neurodegeneration with Axonopathy." *Annual Review of Neuroscience* 35 (1): 153–79. <https://doi.org/10.1146/annurev.neuro.051508.135728>.
- Niemeijer, M., B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, et al. 2010. "Retinopathy Online Challenge: Automatic Detection of Microaneurysms in Digital Color Fundus Photographs." *IEEE Transactions on Medical Imaging* 29 (1): 185–95. <https://doi.org/10.1109/TMI.2009.2033909>.
- Niemeijer, Meindert, Michael D. Abramoff, and Bram van Ginneken. 2009. "Fast Detection of the Optic Disc and Fovea in Color Fundus Photographs." *Medical Image Analysis* 13 (6): 859–70. <https://doi.org/10.1016/j.media.2009.08.003>.
- Niemeijer, Meindert, Bram van Ginneken, Stephen R. Russell, Maria S. A. Suttorp-Schulten, and Michael D. Abramoff. 2007. "Automated Detection and Differentiation of Drusen, Exudates, and Cotton-Wool Spots in Digital Color Fundus Photographs for Diabetic Retinopathy Diagnosis." *Investigative Ophthalmology & Visual Science* 48 (5): 2260–67. <https://doi.org/10.1167/iovs.06-0996>.
- Nishikawa, Robert M. 2007. "Current Status and Future Directions of Computer-Aided Diagnosis in Mammography." *Computerized Medical Imaging and Graphics, Computer-aided Diagnosis (CAD) and Image-guided Decision Support*, 31 (4): 224–35. <https://doi.org/10.1016/j.compmedimag.2007.02.009>.
- Noronha, Kevin P., U. Rajendra Acharya, K. Prabhakar Nayak, Roshan Joy Martis, and Sulatha V. Bhandary. 2014. "Automated Classification of Glaucoma Stages Using

- Higher Order Cumulant Features.” *Biomedical Signal Processing and Control* 10 (March): 174–83. <https://doi.org/10.1016/j.bspc.2013.11.006>.
- Ogurtsova, K., J. D. da Rocha Fernandes, Y. Huang, U. Linnenkamp, L. Guariguata, N. H. Cho, D. Cavan, J. E. Shaw, and L. E. Makaroff. 2017. “IDF Diabetes Atlas: Global Estimates for the Prevalence of Diabetes for 2015 and 2040.” *Diabetes Research and Clinical Practice* 128 (June): 40–50. <https://doi.org/10.1016/j.diabres.2017.03.024>.
- Olmedo, Jesse M., Karen E. Warschaw, Joseph M. Schmitt, and David L. Swanson. 2006. “Optical Coherence Tomography for the Characterization of Basal Cell Carcinoma in Vivo: A Pilot Study.” *Journal of the American Academy of Dermatology* 55 (3): 408–12. <https://doi.org/10.1016/j.jaad.2006.03.013>.
- Ozcan, Aydogan, Alberto Bilenca, Adrien E. Desjardins, Brett E. Bouma, and Guillermo J. Tearney. 2007. “Speckle Reduction in Optical Coherence Tomography Images Using Digital Filtering.” *JOSA A* 24 (7): 1901–10. <https://doi.org/10.1364/JOSAA.24.001901>.
- Pan, S. J., and Q. Yang. 2010. “A Survey on Transfer Learning.” *IEEE Transactions on Knowledge and Data Engineering* 22 (10): 1345–59. <https://doi.org/10.1109/TKDE.2009.191>.
- Pande, Paritosh, Ryan L. Shelton, Guillermo L. Monroy, Ryan M. Nolan, and Stephen A. Boppart. 2017. “Low-Cost Hand-Held Probe for Depth-Resolved Low-Coherence Interferometry.” *Biomedical Optics Express* 8 (1): 338–48. <https://doi.org/10.1364/BOE.8.000338>.
- Paulus, Jan, Jörg Meier, Rüdiger Bock, Joachim Hornegger, and Georg Michelson. 2010. “Automated Quality Assessment of Retinal Fundus Photos.” *International Journal of Computer Assisted Radiology and Surgery* 5 (6): 557–64. <https://doi.org/10.1007/s11548-010-0479-7>.
- Philip, S., A. D. Fleming, K. A. Goatman, S. Fonseca, P. Mcnamee, G. S. Scotland, G. J. Prescott, P. F. Sharp, and J. A. Olson. 2007. “The Efficacy of Automated ‘Disease/No Disease’ Grading for Diabetic Retinopathy in a Systematic Screening Programme.” *British Journal of Ophthalmology* 91 (11): 1512–17. <https://doi.org/10.1136/bjo.2007.119453>.
- Philippakis, E., M. Legrand, M. El Sanharawi, A. Erginay, A. Couturier, and R. Tadayoni. 2018. “Measurement of Full-Thickness Macular Hole Size Using En Face Optical Coherence Tomography.” *Eye (London, England)* 32 (3): 590–96. <https://doi.org/10.1038/eye.2017.254>.
- Piccolo, Domenico, Josef Smolle, Giuseppe Argenziano, Ingrid H. Wolf, Ralph Braun, Lorenzo Cerroni, Angela Ferrari, et al. 2000. “Teledermoscopy - Results of a Multicentre Study on 43 Pigmented Skin Lesions.” *Journal of Telemedicine and Telecare* 6 (3): 132–37. <https://doi.org/10.1258/1357633001935202>.
- Pierre-Filho, Paulo de Tarso Ponte, Rui Barroso Schimiti, José Paulo Cabral de Vasconcellos, and Vital Paulino Costa. 2006. “Sensitivity and Specificity of Frequency-Doubling Technology, Tendency-Oriented Perimetry, SITA Standard and SITA Fast Perimetry in Perimetrically Inexperienced Individuals.” *Acta Ophthalmologica Scandinavica* 84 (3): 345–50. <https://doi.org/10.1111/j.1600-0420.2006.00639.x>.
- Pircher, Michael, Erich Götzinger, Rainer A. Leitgeb, Adolf Friedrich Fercher, and Christoph K. Hitzenberger. 2003a. “Speckle Reduction in Optical Coherence Tomography by Frequency Compounding.” *Journal of Biomedical Optics* 8 (3): 565–70. <https://doi.org/10.1117/1.1578087>.

- Pircher, Michael, Erich Götzinger, Rainer Leitgeb, Adolf F. Fercher, and Christoph K. Hitzenberger. 2003b. "Measurement and Imaging of Water Concentration in Human Cornea with Differential Absorption Optical Coherence Tomography." *Optics Express* 11 (18): 2190–97. <https://doi.org/10.1364/OE.11.002190>.
- Pizurica, A., and W. Philips. 2006. "Estimating the Probability of the Presence of a Signal of Interest in Multiresolution Single- and Multiband Image Denoising." *IEEE Transactions on Image Processing* 15 (3): 654–65. <https://doi.org/10.1109/TIP.2005.863698>.
- Pruthi, Sandhya, Kevin J. Stange, Gerald D. Malagrino, Kashmira S. Chawla, Nicholas F. LaRusso, and Judith S. Kaur. 2013. "Successful Implementation of a Telemedicine-Based Counseling Program for High-Risk Patients With Breast Cancer." *Mayo Clinic Proceedings* 88 (1): 68–73. <https://doi.org/10.1016/j.mayocp.2012.10.015>.
- Puliafito, Carmen A., Michael R. Hee, Charles P. Lin, Elias Reichel, Joel S. Schuman, Jay S. Duker, Joseph A. Izatt, Eric A. Swanson, and James G. Fujimoto. 1995. "Imaging of Macular Diseases with Optical Coherence Tomography." *Ophthalmology* 102 (2): 217–29. [https://doi.org/10.1016/S0161-6420\(95\)31032-9](https://doi.org/10.1016/S0161-6420(95)31032-9).
- Quellec, G., M. Lamard, P. M. Josselin, G. Cazuguel, B. Cochener, and C. Roux. 2008. "Optimal Wavelet Transform for the Detection of Microaneurysms in Retina Photographs." *IEEE Transactions on Medical Imaging* 27 (9): 1230–41. <https://doi.org/10.1109/TMI.2008.920619>.
- Quigley, H. A., and A. T. Broman. 2006. "The Number of People with Glaucoma Worldwide in 2010 and 2020." *British Journal of Ophthalmology* 90 (3): 262–67. <https://doi.org/10.1136/bjo.2005.081224>.
- Quinley, Kelly E., Rachel H. Gormley, Sarah J. Ratcliffe, Ting Shih, Zsofia Szep, Ann Steiner, Doreen Ramogola-Masire, and Carrie L. Kovarik. 2011. "Use of Mobile Telemedicine for Cervical Cancer Screening." *Journal of Telemedicine and Telecare* 17 (4): 203–9. <https://doi.org/10.1258/jtt.2011.101008>.
- Quinn, Graham E., Gui-shuang Ying, Ebenezer Daniel, Lloyd P. Hildebrand, Anna Ells, Agnieszka Baumritter, Alex R. Kemper, Eleanor B. Schron, and Kelly Wade. 2014. "Validity of a Telemedicine System for the Evaluation of Acute-Phase Retinopathy of Prematurity." *Jama Ophthalmology* 132 (10): 1178–84. <https://doi.org/10.1001/jamaophthalmol.2014.1604>.
- Raffle, A. E., B. Alden, M. Quinn, P. J. Babb, and M. T. Brett. 2003. "Outcomes of Screening to Prevent Cancer: Analysis of Cumulative Incidence of Cervical Abnormality and Modelling of Cases and Deaths Prevented." *BMJ* 326 (7395): 901. <https://doi.org/10.1136/bmj.326.7395.901>.
- Raman, Rajiv, Padmaja Kumari Rani, Sheshadri Mahajan, Pradeep Paul, P. Gnanamoorthy, M.s. Krishna, and Tarun Sharma. 2007. "The Tele-Screening Model for Diabetic Retinopathy: Evaluating the Influence of Mydriasis on the Gradability of a Single-Field 45° Digital Fundus Image." *Telemedicine and E-Health* 13 (5): 597–602. <https://doi.org/10.1089/tmj.2006.0084>.
- Rangayyan, Rangaraj M., Fábio J. Ayres, and J. E. Leo Desautels. 2007. "A Review of Computer-Aided Diagnosis of Breast Cancer: Toward the Detection of Subtle Signs." *Journal of the Franklin Institute, Special Issue: Medical Applications of Signal Processing, Part I*, 344 (3): 312–48. <https://doi.org/10.1016/j.jfranklin.2006.09.003>.
- Rasta, Seyed Hossein, Mahsa Eisazadeh Partovi, Hadi Seyedarabi, and Alireza Javadzadeh. 2015. "A Comparative Study on Preprocessing Techniques in Diabetic Retinopathy

- Retinal Images: Illumination Correction and Contrast Enhancement.” *Journal of Medical Signals and Sensors* 5 (1): 40–48.
- Recep, Ömer F., Hikmet Hasırıpı, Nurullah Çağıl, and Hikmet Sarıkatipoğlu. 2001. “Relation between Corneal Thickness and Intraocular Pressure Measurement by Noncontact and Applanation Tonometry.” *Journal of Cataract & Refractive Surgery* 27 (11): 1787–91. [https://doi.org/10.1016/S0886-3350\(01\)00900-2](https://doi.org/10.1016/S0886-3350(01)00900-2).
- Resnikoff, Serge, Donatella Pascolini, Daniel Etya’ale, Ivo Kocur, Ramachandra Pararajasegaram, Gopal P. Pokharel, and Silvio P. Mariotti. 2004. “Global Data on Visual Impairment in the Year 2002.” *Bulletin of the World Health Organization* 82 (November): 844–51. <https://doi.org/10.1590/S0042-96862004001100009>.
- Richter, Grace M., Steven L. Williams, Justin Starren, John T. Flynn, and Michael F. Chiang. 2009. “Telemedicine for Retinopathy of Prematurity Diagnosis: Evaluation and Challenges.” *Survey of Ophthalmology* 54 (6): 671–85. <https://doi.org/10.1016/j.survophthal.2009.02.020>.
- Rogowska, J., and M. E. Brezinski. 2000. “Evaluation of the Adaptive Speckle Suppression Filter for Coronary Optical Coherence Tomography Imaging.” *IEEE Transactions on Medical Imaging* 19 (12): 1261–66. <https://doi.org/10.1109/42.897820>.
- Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, et al. 2015. “ImageNet Large Scale Visual Recognition Challenge.” *International Journal of Computer Vision* 115 (3): 211–52. <https://doi.org/10.1007/s11263-015-0816-y>.
- Saaddine, Jinan B., Amanda A. Honeycutt, K. M. Venkat Narayan, Xinzhi Zhang, Ron Klein, and James P. Boyle. 2008. “Projection of Diabetic Retinopathy and Other Major Eye Diseases Among People With Diabetes Mellitus: United States, 2005-2050.” *Archives of Ophthalmology* 126 (12): 1740–47. <https://doi.org/10.1001/archophth.126.12.1740>.
- Salomon, L. J., Z. Alfrevic, V. Berghella, C. Bilardo, E. Hernandez-Andrade, S. L. Johnsen, K. Kalache, et al. 2011. “Practice Guidelines for Performance of the Routine Mid-Trimester Fetal Ultrasound Scan.” *Ultrasound in Obstetrics & Gynecology* 37 (1): 116–26. <https://doi.org/10.1002/uog.8831>.
- Salvini, Camilla, Daniela Massi, Alessio Cappetti, Marcello Stante, Pietro Cappugi, Paolo Fabbri, and Paolo Carli. 2008. “Application of Optical Coherence Tomography in Non-Invasive Characterization of Skin Vascular Lesions.” *Skin Research and Technology* 14 (1): 89–92. <https://doi.org/10.1111/j.1600-0846.2007.00265.x>.
- Sanchez, C. I., R. Hornero, M. I. Lopez, and J. Poza. 2004. “Retinal Image Analysis to Detect and Quantify Lesions Associated with Diabetic Retinopathy.” In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1:1624–27. <https://doi.org/10.1109/IEMBS.2004.1403492>.
- Sánchez, Clara I., Meindert Niemeijer, Alina V. Dumitrescu, Maria S. A. Suttorp-Schulten, Michael D. Abràmoff, and Bram van Ginneken. 2011. “Evaluation of a Computer-Aided Diagnosis System for Diabetic Retinopathy Screening on Public Data.” *Investigative Ophthalmology & Visual Science* 52 (7): 4866–71. <https://doi.org/10.1167/iovs.10-6633>.
- Sander, B., M. Larsen, L. Thrane, J. L. Hougaard, and T. M. Jørgensen. 2005. “Enhanced Optical Coherence Tomography Imaging by Multiple Scan Averaging.” *British Journal of Ophthalmology* 89 (2): 207–12. <https://doi.org/10.1136/bjo.2004.045989>.

- Sander Dieleman, Jan Schlüter, Colin Raffel, Eben Olson, Søren Kaae Sønderby, Daniel Nouri, Daniel Maturana, et al. 2015. *Lasagne: First Release*. Zenodo. <https://doi.org/10.5281/zenodo.27878>.
- Schmitt, J. M., S. H. Xiang, and K. M. Yung. 1999. "Speckle in Optical Coherence Tomography." *Journal of Biomedical Optics* 4 (1): 95–105. <https://doi.org/10.1117/1.429925>.
- Schmitz, Lutz, Uwe Reinhold, Erhard Bierhoff, and Thomas Dirschka. 2013. "Optical Coherence Tomography: Its Role in Daily Dermatological Practice." *JDDG: Journal Der Deutschen Dermatologischen Gesellschaft* 11 (6): 499–507. <https://doi.org/10.1111/ddg.12073>.
- Schuman, Joel S., Michael R. Hee, Carmen A. Puliafito, Carlton Wong, Tamar Pedut-Kloizman, Charles P. Lin, Ellen Hertzmark, Joseph A. Izatt, Eric A. Swanson, and James G. Fujimoto. 1995. "Quantification of Nerve Fiber Layer Thickness in Normal and Glaucomatous Eyes Using Optical Coherence Tomography: A Pilot Study." *Archives of Ophthalmology* 113 (5): 586–96. <https://doi.org/10.1001/archophth.1995.01100050054031>.
- Sekhar, S., W. Al-Nuaimy, and A. K. Nandi. 2008. "Automated Localisation of Optic Disk and Fovea in Retinal Fundus Images." In *2008 16th European Signal Processing Conference*, 1–5.
- Şenel, Engin, Mete Baba, and Murat Durdu. 2013. "The Contribution of Teledermatoscopy to the Diagnosis and Management of Non-Melanocytic Skin Tumours." *Journal of Telemedicine and Telecare* 19 (1): 60–63. <https://doi.org/10.1177/1357633X12474961>.
- Sevastopolsky, Artem. 2017. "Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network." *Pattern Recognition and Image Analysis* 27 (3): 618–24. <https://doi.org/10.1134/S1054661817030269>.
- Shelton, Ryan L., Woonggyu Jung, Samir I. Sayegh, Daniel T. McCormick, Jeehyun Kim, and Stephen A. Boppart. 2014. "Optical Coherence Tomography for Advanced Screening in the Primary Care Office." *Journal of Biophotonics* 7 (7): 525–33. <https://doi.org/10.1002/jbio.201200243>.
- Shin, H. C., H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers. 2016. "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning." *IEEE Transactions on Medical Imaging* 35 (5): 1285–98. <https://doi.org/10.1109/TMI.2016.2528162>.
- Simonyan, Karen, and Andrew Zisserman. 2014. "Very Deep Convolutional Networks for Large-Scale Image Recognition." *ArXiv:1409.1556 [Cs]*, September. <http://arxiv.org/abs/1409.1556>.
- Sivaswamy, J., S. R. Krishnadas, G. Datt Joshi, M. Jain, and A. U. Syed Tabish. 2014. "Drishti-GS: Retinal Image Dataset for Optic Nerve Head (ONH) Segmentation." In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, 53–56. <https://doi.org/10.1109/ISBI.2014.6867807>.
- Smith, Robert A., Deana Manassaram-Baptiste, Durado Brooks, Mary Doroshenk, Stacey Fedewa, Debbie Saslow, Otis W. Brawley, and Richard Wender. 2015. "Cancer Screening in the United States, 2015: A Review of Current American Cancer Society

- Guidelines and Current Issues in Cancer Screening.” *CA: A Cancer Journal for Clinicians* 65 (1): 30–54. <https://doi.org/10.3322/caac.21261>.
- Soares, J. V. B., J. J. G. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree. 2006. “Retinal Vessel Segmentation Using the 2-D Gabor Wavelet and Supervised Classification.” *IEEE Transactions on Medical Imaging* 25 (9): 1214–22. <https://doi.org/10.1109/TMI.2006.879967>.
- Sopharak, Akara, Bunyarit Uyyanonvara, Sarah Barman, and Thomas H. Williamson. 2008. “Automatic Detection of Diabetic Retinopathy Exudates from Non-Dilated Retinal Images Using Mathematical Morphology Methods.” *Computerized Medical Imaging and Graphics* 32 (8): 720–27. <https://doi.org/10.1016/j.compmedimag.2008.08.009>.
- Srinivasan, Vivek J., Sava Sakadžić, Iwona Gorczynska, Svetlana Ruvinskaya, Weicheng Wu, James G. Fujimoto, and David A. Boas. 2010. “Quantitative Cerebral Blood Flow with Optical Coherence Tomography.” *Optics Express* 18 (3): 2477–94. <https://doi.org/10.1364/OE.18.002477>.
- Srinivasan, Vivek J., Maciej Wojtkowski, Andre J. Witkin, Jay S. Duker, Tony H. Ko, Mariana Carvalho, Joel S. Schuman, Andrzej Kowalczyk, and James G. Fujimoto. 2006. “High-Definition and 3-Dimensional Imaging of Macular Pathologies with High-Speed Ultrahigh-Resolution Optical Coherence Tomography.” *Ophthalmology* 113 (11): 2054–2065.e3. <https://doi.org/10.1016/j.ophtha.2006.05.046>.
- Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting.” *J. Mach. Learn. Res.* 15 (1): 1929–1958.
- Stahl, A., M. Boeker, C. Ehlken, H. Agostini, and T. Reinhard. 2009. “Evaluation of an internet-based e-learning ophthalmology module for medical students.” *Der Ophthalmologe: Zeitschrift der Deutschen Ophthalmologischen Gesellschaft* 106 (11): 999–1005. <https://doi.org/10.1007/s00347-009-1916-2>.
- Stang, Andreas, and Karl-Heinz Jöckel. 2016. “Does Skin Cancer Screening Save Lives? A Detailed Analysis of Mortality Time Trends in Schleswig-Holstein and Germany.” *Cancer* 122 (3): 432–37. <https://doi.org/10.1002/cncr.29755>.
- Stiffler, Meghan C., and Bridget V. Dever. 2015. “History of Screening Practices, Mental Health Assessment, and Classification in the USA.” In *Mental Health Screening at School*, 5–26. Contemporary Issues in Psychological Assessment. Springer, Cham. https://doi.org/10.1007/978-3-319-19171-3_2.
- Swanson, E. A., J. A. Izatt, M. R. Hee, D. Huang, C. P. Lin, J. S. Schuman, C. A. Puliafito, and J. G. Fujimoto. 1993. “In Vivo Retinal Imaging by Optical Coherence Tomography.” *Optics Letters* 18 (21): 1864–66. <https://doi.org/10.1364/OL.18.001864>.
- Swensen, Stephen J. 2003. “Screening for Cancer with Computed Tomography.” *BMJ: British Medical Journal* 326 (7395): 894–95.
- Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. “Going Deeper With Convolutions.” In , 1–9. https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deeper_With_2015_CVPR_paper.html.
- Szkulmowski, Maciej, Iwona Gorczynska, Daniel Szlag, Marcin Sylwestrzak, Andrzej Kowalczyk, and Maciej Wojtkowski. 2012. “Efficient Reduction of Speckle Noise in

- Optical Coherence Tomography.” *Optics Express* 20 (2): 1337–59. <https://doi.org/10.1364/OE.20.001337>.
- T Gambichler, J Hyun, G Moussa, N S Tomi, S Boms, P Altmeyer, K Hoffmann, and A Kreuter. 2007. “Optical Coherence Tomography of Cutaneous Lupus Erythematosus Correlates with Histopathology.” *Lupus* 16 (1): 35–38. <https://doi.org/10.1177/0961203306074773>.
- Tajbakhsh, N., J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. 2016. “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?” *IEEE Transactions on Medical Imaging* 35 (5): 1299–1312. <https://doi.org/10.1109/TMI.2016.2535302>.
- Takizawa, M., S. Sone, K. Hanamura, and K. Asakura. 2001. “Telemedicine System Using Computed Tomography van of High-Speed Telecommunication Vehicle.” *IEEE Transactions on Information Technology in Biomedicine* 5 (1): 2–9. <https://doi.org/10.1109/4233.908348>.
- Tamura, Hiroshi, Rei Goto, Yoko Akune, Yoshimune Hiratsuka, Shusuke Hiragi, and Masakazu Yamada. 2015. “The Clinical Effectiveness and Cost-Effectiveness of Screening for Age-Related Macular Degeneration in Japan: A Markov Modeling Study.” *PLOS ONE* 10 (7): e0133628. <https://doi.org/10.1371/journal.pone.0133628>.
- Tan, Jen Hong, U. Rajendra Acharya, Sulatha V. Bhandary, Kuang Chua Chua, and Sobha Sivaprasad. 2017. “Segmentation of Optic Disc, Fovea and Retinal Vasculature Using a Single Convolutional Neural Network.” *Journal of Computational Science* 20 (May): 70–79. <https://doi.org/10.1016/j.jocs.2017.02.006>.
- Tham, Yih-Chung, Xiang Li, Tien Y. Wong, Harry A. Quigley, Tin Aung, and Ching-Yu Cheng. 2014. “Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040: A Systematic Review and Meta-Analysis.” *Ophthalmology* 121 (11): 2081–90. <https://doi.org/10.1016/j.ophtha.2014.05.013>.
- “The National Lung Screening Trial: Overview and Study Design.” 2011. *Radiology* 258 (1): 243–53. <https://doi.org/10.1148/radiol.10091808>.
- Thomas, Sera-Melisa, Maya Jeyaraman, William G. Hodge, Cindy Hutnik, John Costella, and Monali S. Malvankar-Mehta. 2014. “The Effectiveness of Teleglaucoma versus In-Patient Examination for Glaucoma Screening: A Systematic Review and Meta-Analysis.” *PLOS ONE* 9 (12): e113779. <https://doi.org/10.1371/journal.pone.0113779>.
- Thorner, R. M., and Q. R. Remein. 1961. “Principles and Procedures in the Evaluation of Screening for Disease.” *Principles and Procedures in the Evaluation of Screening for Disease.*, no. 846. <https://www.cabdirect.org/cabdirect/abstract/19662700564>.
- Tielsch, James M., Joanne Katz, Kuldev Singh, Harry A. Quigley, John D. Gottsch, Jonathan Javitt, and Alfred Sommer. 1991. “A Population-Based Evaluation of Glaucoma Screening: The Baltimore Eye Survey.” *American Journal of Epidemiology* 134 (10): 1102–10. <https://doi.org/10.1093/oxfordjournals.aje.a116013>.
- Ting, Daniel Shu Wei, Gemmy Chui Ming Cheung, and Tien Yin Wong. 2016. “Diabetic Retinopathy: Global Prevalence, Major Risk Factors, Screening Practices and Public Health Challenges: A Review.” *Clinical & Experimental Ophthalmology* 44 (4): 260–77. <https://doi.org/10.1111/ceo.12696>.
- Tuulonen, Anja. 2011. “Economic Considerations of the Diagnosis and Management for Glaucoma in the Developed World.” *Current Opinion in Ophthalmology* 22 (2): 102. <https://doi.org/10.1097/ICU.0b013e3283437bf0>.

- Ulrich, M., E. Stockfleth, J. Roewert-Huber, and S. Astner. 2007. "Noninvasive Diagnostic Tools for Nonmelanoma Skin Cancer." *British Journal of Dermatology* 157 (December): 56–58. <https://doi.org/10.1111/j.1365-2133.2007.08275.x>.
- Waldmann, A., S. Nolte, M. A. Weinstock, E. W. Breitbart, N. Eisemann, A. C. Geller, R. Greinert, B. Volkmer, and A. Katalinic. 2012. "Skin Cancer Screening Participation and Impact on Melanoma Incidence in Germany – an Observational Study on Incidence Trends in Regions with and without Population-Based Screening." *British Journal of Cancer* 106 (5): 970–74. <https://doi.org/10.1038/bjc.2012.22>.
- Walter, Thomas, and Jean-Claude Klein. 2001. "Segmentation of Color Fundus Images of the Human Retina: Detection of the Optic Disc and the Vascular Tree Using Morphological Techniques." In *Medical Data Analysis*, 282–87. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45497-7_43.
- Warshaw, Erin M., Yonatan J. Hillman, Nancy L. Greer, Emily M. Hagel, Roderick MacDonald, Indulis R. Rutks, and Timothy J. Wilt. 2011. "Teledermatology for Diagnosis and Management of Skin Conditions: A Systematic Review." *Journal of the American Academy of Dermatology* 64 (4): 759–772.e21. <https://doi.org/10.1016/j.jaad.2010.08.026>.
- Warshaw, Erin M., Frank A. Lederle, Joseph P. Grill, Amy A. Gravely, Ann K. Bangerter, Lawrence A. Fortier, Kimberly A. Bohjanen, et al. 2009. "Accuracy of Teledermatology for Nonpigmented Neoplasms." *Journal of the American Academy of Dermatology* 60 (4): 579–88. <https://doi.org/10.1016/j.jaad.2008.11.892>.
- Weinreb, Robert N., Tin Aung, and Felipe A. Medeiros. 2014. "The Pathophysiology and Treatment of Glaucoma: A Review." *JAMA* 311 (18): 1901–11. <https://doi.org/10.1001/jama.2014.3192>.
- Weinreb, Robert N., Linda Zangwill, Charles C. Berry, Renuka Bathija, and Pamela A. Sample. 1998. "Detection of Glaucoma With Scanning Laser Polarimetry." *Archives of Ophthalmology* 116 (12): 1583–89. <https://doi.org/10.1001/archophth.116.12.1583>.
- Wells, F. L. 1920. Review of *Review of Army Mental Tests*, by C. S. Yoakum and R. M. Yerkes. *The Journal of Philosophy, Psychology and Scientific Methods* 17 (11): 304–6. <https://doi.org/10.2307/2940600>.
- Welzel, Julia, Maike Bruhns, and Helmut H. Wolff. 2003. "Optical Coherence Tomography in Contact Dermatitis and Psoriasis." *Archives of Dermatological Research* 295 (2): 50–55. <https://doi.org/10.1007/s00403-003-0390-y>.
- Weng, Sheng, Xiaoyun Xu, Jiasong Li, and Stephen T. C. Wong. 2017. "Combining Deep Learning and Coherent Anti-Stokes Raman Scattering Imaging for Automated Differential Diagnosis of Lung Cancer." *Journal of Biomedical Optics* 22 (10): 106017. <https://doi.org/10.1117/1.JBO.22.10.106017>.
- WHO. 2010. "Telemedicine. Opportunities and Developments in Member States. Report on the Second Survey on EHealth. Global Observatory for EHealth Series - Volume 2." 2010. http://www.who.int/goe/publications/goe_telemedicine_2010.pdf.
- Wilkerson, Hugh L. C., and Leo P. Krall. 1947. "Diabetes in a New England Town: A Study of 3,516 Persons in Oxford, Mass." *Journal of the American Medical Association* 135 (4): 209–16. <https://doi.org/10.1001/jama.1947.02890040011003>.
- Wilson, J. M. G., and G. Jungner. 1968. "Principles and Practice of Screening for Disease."

- Winetraub, Yonatan, Chris Wu, Steven Chu, and Adam de la Zerda. 2017. "Upper Limit for Angular Compounding Speckle Reduction." *BioRxiv*, December, 239350. <https://doi.org/10.1101/239350>.
- Wolff, Tracy. 2009. "Screening for Skin Cancer: An Update of the Evidence for the U.S. Preventive Services Task Force." *Annals of Internal Medicine* 150 (3): 194. <https://doi.org/10.7326/0003-4819-150-3-200902030-00009>.
- Wong, D. W. K., J. Liu, J. H. Lim, X. Jia, F. Yin, H. Li, and T. Y. Wong. 2008. "Level-Set Based Automatic Cup-to-Disc Ratio Determination Using Retinal Fundus Images in ARGALI." In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2266–69. <https://doi.org/10.1109/IEMBS.2008.4649648>.
- Wong, Hon-Tym, Marcus C. Lim, Lisandro M. Sakata, Han T. Aung, Nishani Amerasinghe, David S. Friedman, and Tin Aung. 2009. "High-Definition Optical Coherence Tomography Imaging of the Iridocorneal Angle of the Eye." *Archives of Ophthalmology* 127 (3): 256–60. <https://doi.org/10.1001/archophthalmol.2009.22>.
- Wong, Wan Ling, Xinyi Su, Xiang Li, Chui Ming G Cheung, Ronald Klein, Ching-Yu Cheng, and Tien Yin Wong. 2014. "Global Prevalence of Age-Related Macular Degeneration and Disease Burden Projection for 2020 and 2040: A Systematic Review and Meta-Analysis." *The Lancet Global Health* 2 (2): e106–16. [https://doi.org/10.1016/S2214-109X\(13\)70145-1](https://doi.org/10.1016/S2214-109X(13)70145-1).
- Xu, Tao, Han Zhang, Xiaolei Huang, Shaoting Zhang, and Dimitris N. Metaxas. 2016. "Multimodal Deep Learning for Cervical Dysplasia Diagnosis." In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, 115–23. Lecture Notes in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-46723-8_14.
- Yamada, N., P. P. Chen, R. P. Mills, M. M. Leen, R. L. Stamper, M. F. Lieberman, L. Xu, and D. C. Stanford. 2000. "Glaucoma Screening Using the Scanning Laser Polarimeter." *Journal of Glaucoma* 9 (3): 254–61.
- Yogesana, K., I. J. Constable, C. J. Barry, R. H. Eikelboom, W. Morgan, M. L. Tay-Kearney, and L. Jitskaia. 1999. "Evaluation of a Portable Fundus Camera for Use in the Teleophthalmologic Diagnosis of Glaucoma." *Journal of Glaucoma* 8 (5): 297–301.
- Yogesana, Kanagasingam, Ian J. Constable, Chris J. Barry, Robert H. Eikelboom, Ian L. McAllister, and Mei-Ling Tay-Kearney. 2000. "Telemedicine Screening of Diabetic Retinopathy Using a Hand-Held Fundus Camera." *Telemedicine Journal* 6 (2): 219–23. <https://doi.org/10.1089/107830200415153>.
- Yogesana, Kanagasingam, Sajeesh Kumar, Leonard Goldschmidt, and Jorge Cuadros. 2006. *Teleophthalmology*. Springer Science & Business Media.
- Zaki, Farzana, Yahui Wang, Hao Su, Xin Yuan, and Xuan Liu. 2017. "Noise Adaptive Wavelet Thresholding for Speckle Noise Removal in Optical Coherence Tomography." *Biomedical Optics Express* 8 (5): 2720–31. <https://doi.org/10.1364/BOE.8.002720>.
- Zalaudek, Iris, Giuseppe Argenziano, Alessandro Di Stefani, Gerardo Ferrara, Ashfaq A. Marghoob, Rainer Hofmann-Wellenhof, H. Peter Soyer, Ralph Braun, and Helmut Kerl. 2006. "Dermoscopy in General Dermatology." *Dermatology* 212 (1): 7–18. <https://doi.org/10.1159/000089015>.
- Zhang, Bob, Xiangqian Wu, Jane You, Qin Li, and Fakhri Karray. 2009. "Hierarchical Detection of Red Lesions in Retinal Images by Multiscale Correlation Filtering." In

- Medical Imaging 2009: Computer-Aided Diagnosis*, 7260:72601L. International Society for Optics and Photonics. <https://doi.org/10.1117/12.813913>.
- Zhao, Di, Eliseo Guallar, Prateek Gajwani, Bonnielin Swenor, John Crews, Jinan Saaddine, Lucy Mudie, et al. 2017. "Optimizing Glaucoma Screening in High-Risk Population: Design and 1-Year Findings of the Screening to Prevent (SToP) Glaucoma Study." *American Journal of Ophthalmology* 180 (August): 18–28. <https://doi.org/10.1016/j.ajo.2017.05.017>.
- Zhao, Yingru, Geertruida H. de Bock, Rozemarijn Vliegenthart, Rob J. van Klaveren, Ying Wang, Luca Bogoni, Pim A. de Jong, Willem P. Mali, Peter M. A. van Ooijen, and Matthijs Oudkerk. 2012. "Performance of Computer-Aided Detection of Pulmonary Nodules in Low-Dose CT: Comparison with Double Reading by Nodule Volume." *European Radiology* 22 (10): 2076–84. <https://doi.org/10.1007/s00330-012-2437-y>.
- Zheng, Yingfeng, Mingguang He, and Nathan Congdon. 2012. "The Worldwide Epidemic of Diabetic Retinopathy." *Indian Journal of Ophthalmology* 60 (5): 428–31. <https://doi.org/10.4103/0301-4738.100542>.
- Zilly, Julian, Joachim M. Buhmann, and Dwarikanath Mahapatra. 2017. "Glaucoma Detection Using Entropy Sampling and Ensemble Learning for Automatic Optic Cup and Disc Segmentation." *Computerized Medical Imaging and Graphics*, Special Issue on Ophthalmic Medical Image Analysis, 55 (January): 28–41. <https://doi.org/10.1016/j.compmedimag.2016.07.012>.

Journal articles

- Guerra, P., Valverde, J.J., Martin, A., Ledesma, M.J., Rubio-Guivernau, J.L., Santos, A. "Real Time Signal Processing and Data Handling with dedicated hardware in handheld OCT Device". J. Instrum., 10(11):C11001. Nov. 2015 (doi: 10.1088/1748-0221/10/11/C11001).
- Juan J. Gómez-Valverde, Zhe Chen, Christoph Sinz, Andres Santos, Wolfgang Drexler, María J. Ledesma-Carbayo. "Wavelet compounding adaptive speckle noise reduction filter for optical coherence tomography skin imaging". Manuscript to be submitted to Biomedical Optic Express.
- Juan J. Gómez-Valverde, Alfonso Antón, Gianluca Fatti, Andrés Santos, Clara I. Sánchez, María J. Ledesma-Carbayo. "Automatic Glaucoma Classification of color fundus images based on convolutional neural networks and transfer learning". Manuscript to be submitted to Biomedical Optic Express.

Conference Proceedings

- Gómez-Valverde, J.J., Ortuño J.E., Guerra P., Hermann B., Zabihian B., Rubio-Guivernau J.L., Santos A., Drexler W., Ledesma-Carbayo M.J. "Evaluation of speckle reduction with denoising filtering in optical coherence tomography for dermatology". IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2015), pp. 494-497. New York (USA), Apr. 2015.
- Gómez-Valverde, J.J., Ortuño J.E., Guerra P., Hermann B., Zabihian B., Rubio-Guivernau J.L., Santos A., Drexler W., Ledesma-Carbayo M.J. "Speckle reduction process based on digital filtering and wavelet compounding in optical coherence tomography for dermatology". Proceedings of SPIE-OSA Biomedical Optics Vol. 9541, pp. 95411K, Munich, June 2015
- Gomez, J.J., Sanchez, C.I., Liefers, B., Venhuizen, F.G., Fatti, G., Morilla-Grasa, A., Cartagena, Y., Herranz Cabarcos, A., Santos, A., Ledesma-Carbayo, M.J., Anton-Lopez, A. "Automated Analysis of Retinal Images for the detection of Glaucoma based on Convolutional Neural Networks". Invest. Ophthalmol. Vis. Sci., 58(8):707. (ARVO 2017: The annual meeting of the Association for Research in Vision and Ophthalmology). Jun. 2017.
- Alfonso Anton, Gianluca Fatty, Alejandra Herranz, Eleonora Ayala, Antonio Morilla, Juan Jose Gomez Valverde. "DETECTION OF GLAUCOMA WITH TELEMEDICINE OPTIC DISK PHOTOS AND OCT IN A PRIMARY CARE CENTRE". 13th European Glaucoma Society Congress. Florence (Italy). May 2018.

- Alfonso Antón, Gianluca Fatti, Alejandra Herranz, Eleonora Ayala, Isaac Alarcón, Antonio Morilla, Yasmín, Zaida and Juan J. Gómez. “Detection of Glaucoma, Diabetic Retinopathy and Age Related Macular Degeneration with Telemedicine, Photos and OCT”. World Ophthalmology Congress (WOC 2018). Barcelona. June 2018.
- J.J. Gómez-Valverde, A. Antón-López, G. Fatti, M.J. Ledesma-Carbayo, A. Santos-Lleó. “Glaucoma screening based on multimodal image-guided diagnosis, machine learning and telemedicine”. 15th meeting of the Imaging and Morphometry Association for Glaucoma in Europe (IMAGE). March 2016. Madrid.
- J.J. Gómez-Valverde, C.I. Sánchez, G. Fatti, M.J. Ledesma-Carbayo, A. Santos-Lleó, A. Antón-López 16th “Automated Analysis of Retinal Images for detection of Glaucoma based on Convolutional Neural Networks” meeting of the Imaging and Morphometry Association for Glaucoma in Europe (IMAGE). March 2017. Rome (Italy)
- J. Gómez Valverde, G. Fatti, A. Antón López. “Rastreo de glaucoma basado en detección guiada por imágenes, análisis automático de imágenes y telemedicina”. XXXIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, pág. 426-429.. Noviembre 2015. Madrid.
- P. Guerra Gutierrez, J.J. Gómez Valverde, A. Martín Gómez, M.J. Ledesma Carbayo, J.L. Rubio Guivernau, A. Santos Lleó. “Multichannel quadrature coherent demodulator for a portable time-domain OCT device”. XXXIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, pág. 467-470. Noviembre 2015. Madrid

Software Registry

- Software registry in the Registry of Intellectual Property of Comunidad de Madrid (Spain). Telemedicine platform “DYSEO, Sistema de Detección y Seguimiento de Enfermedades Oculares”. May 2016.

Agradecimientos

Quiero finalizar dando las gracias a todas las personas que me han ayudado estos años durante la realización de la tesis doctoral. Primero a mis directores, María Jesús Ledesma y Andrés Santos. Siempre les agradeceré su paciencia, flexibilidad, soporte profesional y en especial su cariñoso apoyo personal en momentos difíciles. Su competencia y gran talento profesional han sido clave para poder concluir con éxito muchos de los objetivos que nos marcamos al iniciar el trabajo.

Sin duda una de las experiencias profesionales y personales más enriquecedoras de la tesis ha sido tener la posibilidad de viajar y trabajar con otros grupos de investigación. He tenido la suerte de colaborar con equipos de primer nivel internacional en un ambiente cercano que me ha hecho sentir como en casa. Quiero agradecer especialmente a Clarisa Sánchez por permitirme trabajar en su equipo como uno más y por su guía, supervisión y apoyo durante mi estancia en Nimega. Agradezco a su equipo, especialmente a Bart Liefers y Freerk Venhuizen, su ayuda y apoyo técnico en todo momento. También quiero agradecer al profesor Wolfgang Drexler por su amable acogida y soporte durante mi estancia en Viena. Poder trabajar junto a su equipo ha sido un privilegio del que me siento muy afortunado. No puedo dejar de dar las gracias a Zhe Chen, Elisabet Rank y a Christoph Sinz por su ayuda y generosidad durante el tiempo que estuve allí. Durante la realización de la tesis he tenido la suerte de poder trabajar en proyectos de investigación apasionantes con grandes profesionales de los que he aprendido mucho. Quiero agradecer a José Luis Rubio y a Pedro Guerra por haberme permitido conocer de cerca los enormes retos a los que se enfrenta una empresa de invocación tecnológica y haber compartido conmigo su experiencia y grandes conocimientos.

Hubiera sido imposible para mí completar este trabajo sin la ayuda de mi familia. Siempre les agradeceré su incondicional y total apoyo todos estos años. Mi madre, Tere, un ejemplo constante de generosidad y esfuerzo, Alicia y Virginia, las mejores y más cariñosas hermanas que se puedan tener, Ana, mi pareja y compañera de vida pero también de congresos y estancias, siempre comprensiva y cariñosa. Y mis hijos, David y Raquel, que nacieron al empezar la tesis y me han dado las mayores alegrías desde entonces. A todos ellos gracias y todo mi amor.

Me gustaría concluir teniendo un especial recuerdo para mi padre. Siento no poder compartir con él la satisfacción de acabar esta tesis doctoral de la que tan orgulloso y feliz se habría sentido. Su pasión por el trabajo, su honestidad, ética personal y amor por su familia han sido para mí y serán siempre una inspiración y referencia constante. Siempre me acompañara tu recuerdo.