

# Tesis Doctoral

2018



**Detección y Segmentación Simultánea de  
Estructuras Anatómicas en Imágenes de  
Retina basada en Modelos de Conocimiento  
Relacionales Intra e Inter-Estructura**

**José María Molina Casado**

**Programa de Doctorado en Sistemas Inteligentes**

**Director: Enrique J. Carmona Suárez**



UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

Departamento de Inteligencia Artificial

*Escuela Técnica Superior de Ingeniería Informática*



# **Detección y Segmentación Simultánea de Estructuras Anatómicas en Imágenes de Retina basada en Modelos de Conocimiento Relacionales Intra e Inter-Estructura**

**José María Molina Casado**

Ingeniero Técnico en Informática de Gestión por la Universidad de Jaén

Ingeniero en Informática por la Universidad de Granada

Máster en Inteligencia Artificial Avanzada por la UNED

Director

**Enrique J. Carmona Suárez**

Profesor Titular en el Departamento de Inteligencia Artificial

en la Universidad Nacional de Educación a Distancia



©2018 José María Molina Casado  
Este documento ha sido generado  
utilizando *Lyx*, L<sup>A</sup>T<sub>E</sub>X y otras  
herramientas de open-source.



*“El saber no ocupa lugar... pero sí tiempo”*

Anónimo



*A mis padres, Lázaro y Alcázar*



# Agradecimientos

En primer lugar quiero agradecer sinceramente a mi director de tesis Enrique Carmona su paciencia conmigo en todo momento; por hacerme ver que siempre hay que analizar y cuestionarlo todo; por enseñarme a no lanzar las campanas al vuelo ante un buen resultado y, al contrario, a no bajar los brazos ante uno malo. Finalizando la ardua tarea de elaborar una tesis doctoral me doy cuenta que, mucho más importante que los conocimientos adquiridos, son las sensaciones que subyacen y se aprenden por el camino. Ahora, sin duda, me siento aún más paciente y crítico con mi trabajo. Todo esto es gracias a él.

En segundo lugar, dar las gracias a mi hermano, a mis tíos y tías, a mis abuelos que ya no están... Pero principalmente dar gracias a mis padres por haberme educado y también por darme una educación, porque aunque lo parezca, no significa lo mismo. Por un lado soy como soy gracias a ellos, por esa educación de base aprendida en casa, mucho más importante que cualquier título académico. Por otro lado me han facilitado la posibilidad de estudiar y formarme, apoyándome sin descanso y sin dejar de creer en mí en todo momento. Por todo y mucho más, a todos, os quiero.

En tercer lugar, dar las gracias a toda mi gente de Cazorla, a aquellos que tanto me han escuchado hablar de *los ojos* (cariñoso apodo con el que me refiero a la tesis desde hace años). Gracias por perdonar que no asistiera a veces a la cerveza de turno por terminar mis tareas pendientes. Os aseguro que las recuperaremos retroactivamente!

Y en último lugar, y por supuesto no menos importante, agradecer a mi Alicia todo su apoyo constante y comprensión infinita. Sus ánimos, su paciencia, su todo. Por ser la mejor persona que conozco. Eres mi principal motivación e intento superarme y ser mejor persona gracias a ti. Te quiero.

Porque sin vosotros no hubiera tenido ni la fuerza ni la motivación necesarias para hacer este trabajo, a todos y cada uno, gracias.



# Resumen

La detección y segmentación de estructuras anatómicas en imágenes de retina constituyen dos tareas previas de gran importancia para la implementación de cualquier sistema automático de diagnóstico de patologías retinianas, siendo hoy día dos problemas aún abiertos. Por ejemplo, los resultados de detección de las diferentes estructuras anatómicas pueden utilizarse como un paso previo a la segmentación o como puntos de referencia usados por sistemas automáticos para el diagnóstico de diferentes patologías retinianas. Por otra parte, la segmentación del disco óptico es fundamental, por ejemplo, para el diagnóstico de enfermedades como el glaucoma o la retinopatía diabética proliferativa. Igualmente, la segmentación de la fovea permite, por ejemplo, graduar la gravedad de determinadas maculopatías, permitiendo calcular la distancia de determinados patrones de patología al centro de esta estructura.

La mayoría de los trabajos en la literatura relacionada han estado orientados a la detección/segmentación de cada estructura individualizada, o a la detección/segmentación de dos o más estructuras realizada de forma secuencial. Por el contrario, en esta tesis, se proponen dos metodologías, una de detección y otra de segmentación, que comparten tres propiedades importantes: (i) son independientes del tamaño de la imagen de entrada; (ii) están orientadas al uso conjunto del conocimiento relacional asociado a cada estructura retiniana (CR<sub>intra-e</sub>) y al de varias de estas estructuras (CR<sub>inter-e</sub>); y (iii) el uso conjunto de ambos tipos de conocimiento permite realizar la detección/segmentación de  $n$  estructuras de forma simultánea. La primera característica persigue que la configuración de parámetros de cualquier método basado en estas metodologías sea independiente del tamaño de imagen y, la segunda y tercera, que el número de falsos positivos disminuya respecto de aquellos casos en que la

detección o segmentación se realiza de forma individual o secuencial.

A partir de las dos metodologías mencionadas anteriormente, se detallan los pasos seguidos para la implementación de sendos métodos de detección y segmentación de estructuras retinianas. Uno de estos pasos hace referencia a la construcción de una serie de modelos relacionales de cada una de las estructuras a detectar/segmentar, distinguiendo entre modelos que solo usan información de la propia estructura (modelos intra-e) o de más de una estructura (modelos inter-e). De esta forma, se construyen cuatro modelos intra-e (disco óptico, mácula, red de vasos principales y haz vascular) y un modelo implícito inter-e (basado en restricciones) que son usados por el primer método para la detección simultánea de las cuatro estructuras mencionadas. De otro lado, se construyen dos modelos intra-e (disco óptico y fóvea) y un modelo relacional para ser usado por el segundo método para la segmentación simultánea de las dos estructuras mencionadas. La principal diferencia entre los modelos intra-e usados para detección y segmentación radica en su complejidad, siendo muy simples en el primer caso y algo más complejos en el segundo.

Los resultados de los experimentos diseñados refrendan las bondades de la simultaneidad y el uso conjunto del CR<sub>intra</sub> e <sub>inter-e</sub>, tanto en detección como en segmentación, siendo la disminución de la tasa de falsos positivos una de sus principales consecuencias. Todos los experimentos se realizaron en bases de imágenes públicas para facilitar la comparación con otros métodos de detección/segmentación. Un experimento adicional muestra que la incorporación del resultado de detección como entrada adicional al método de segmentación favorece los resultados de este último, a costa de un pequeño incremento del tiempo empleado en el proceso.

# Summary

Detection and segmentation of anatomical structures in retinal images are two previously required tasks for the implementation of any automatic system oriented to the diagnosis of retinal pathologies. Despite the fact of their outstanding importance, both tasks remain open problems to this day. As an example, results of the detection for the different anatomic structures can be used as a prior step to segmentation, or even as landmarks for automatized diagnosis systems with the purpose of detecting any illness symptomatology. On the other hand, optic disk segmentation is of capital importance to detect diseases as glaucoma or proliferative diabetic retinopathy. Equally, fovea segmentation enables the measurement of the distance to possible disease patterns, evaluating the grade of certain maculopathies.

Most of the literary works related to this subject have been developed in two different ways: focusing in the detection/segmentation of each structure as an individualized item or carrying out the detection/segmentation of two or more structures in a sequential way. Conversely, this thesis proposes two methodologies, one for detection and another for segmentation, sharing both of them three relevant features: (i) they are independent of the size of the input image; (ii) they are oriented to the joint use of the relational knowledge associated with each retinal structure (intra-SRK) and with other structures (inter-SRK); and (iii) the joint use of both types of knowledge allows the execution of the detection/segmentation of  $n$  structures simultaneously. The first characteristic intends the parameter configuration of any method based in these methodologies to be independent from the image size. In addition, the second and third features are aimed to diminish the number of false positives in relation to those cases in which just one structure is detected/segmented or detection/segmentation

is sequentially implemented.

On the basis of the two above mentioned methodologies, the necessary steps to implement a method for retinal structures detection/segmentation are described. One of these steps is associated with the construction of relational models for each one of the structures destined to be detected/segmented, distinguishing between models that employ knowledge from just one structure (intra-SRK models) and those that use knowledge from more than one structure (inter-SRK model). This way, four intra-SRK models are built (being related to optic disk, macula, main retinal vessel network and vascular bundle), plus an implicit inter-SRK model (based on restrictions). All of them are used by the first method to simultaneously detect the four above mentioned structures. On the other side, two intra-SRK models (optic disk and fovea) and a relational model are used by the second method for simultaneously segment the two structures previously referred. The main difference between intra-SRK models used for detection and segmentation lies in its degree of complexity, resulting in relatively simple models in the first case, but somewhat more complex in the second.

The results from the designed experiments show the benefits of simultaneity property and the joint use of intra-SRK and inter-SRK for both detection and segmentation, being the decrease of the false positive rate its main consequence. All experiments were carried out on public databases to make easier the comparison to other detection/segmentation methods. An additional experiment shows that the incorporation of the detection result as an additional input for the segmentation method favors the results obtained by this last one, at the expense of a little increase in the computational time spent to complete the process.

# Índice general

<b>Resumen</b>	<b>I</b>
<b>Summary</b>	<b>III</b>
<b>Lista de acrónimos</b>	<b>XIII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Alcance de la tesis . . . . .	5
1.2.1. Descripción del problema . . . . .	5
1.2.2. Limitaciones . . . . .	10
1.2.3. Hipótesis y objetivos de investigación . . . . .	10
1.3. Materiales . . . . .	13
1.4. Estructura de la tesis . . . . .	14
<b>2. Estado del arte</b>	<b>17</b>
2.1. Preprocesado en imágenes de retina . . . . .	17
2.2. Detección de estructuras anatómicas retinianas . . . . .	20
2.2.1. Uso aislado del CRintra-e . . . . .	21
2.2.2. Uso conjunto de CRintra e inter-e . . . . .	24
2.3. Segmentación de estructuras anatómicas retinianas . . . . .	40
2.3.1. Uso aislado de CRintra-e . . . . .	40
2.3.2. Uso conjunto de CRintra e inter-e . . . . .	47

<b>3. Metodología</b>	<b>59</b>
3.1. Detección de estructuras retinianas . . . . .	60
3.1.1. Preprocesado . . . . .	60
3.1.2. Utilizando conocimiento relacional intra-estructura . . . . .	63
3.1.3. Utilizando conocimiento relacional inter-estructura . . . . .	71
3.2. Segmentación de estructuras retinianas . . . . .	78
3.2.1. Preprocesado . . . . .	79
3.2.2. Modelo de Disco Óptico . . . . .	80
3.2.3. Modelo de Fóvea . . . . .	87
3.2.4. Modelo relacional DO-Fóvea . . . . .	91
3.2.5. Uso de Algoritmos Evolutivos . . . . .	93
3.3. Segmentación con información de detección . . . . .	98
<b>4. Resultados experimentales y discusión</b>	<b>103</b>
4.1. Diseño de experimentos de detección . . . . .	104
4.2. Resultados de detección . . . . .	106
4.2.1. Evaluación en la base de imágenes MESSIDOR . . . . .	106
4.2.2. Evaluación en otras bases de imágenes públicas . . . . .	109
4.3. Discusión de resultados de detección . . . . .	113
4.3.1. Base de imágenes MESSIDOR . . . . .	113
4.3.2. Otras bases de imágenes . . . . .	115
4.3.3. Otras consideraciones . . . . .	118
4.4. Diseño de experimentos de segmentación . . . . .	123
4.5. Resultados de segmentación . . . . .	127
4.5.1. Evaluación en la base de imágenes MESSIDOR . . . . .	127
4.5.2. Evaluación en la base de imágenes ONHSD . . . . .	131
4.6. Discusión de resultados de segmentación . . . . .	135
4.6.1. Base de imágenes MESSIDOR . . . . .	135
4.6.2. Base de imágenes ONHSD . . . . .	136

4.6.3. Coste computacional . . . . .	137
4.7. Resultados de segmentación con información de detección . . . . .	138
<b>5. Conclusiones y trabajo futuro</b>	<b>141</b>
5.1. Conclusiones . . . . .	141
5.2. Trabajo futuro . . . . .	143
<b>A. Publicaciones asociadas a la tesis</b>	<b>145</b>
<b>B. Conclusions and future work</b>	<b>147</b>
B.1. Conclusions . . . . .	147
B.2. Future work . . . . .	149
<b>Bibliografía</b>	<b>150</b>



# Índice de figuras

1.1. Principales estructuras anatómicas a considerar en una imagen de retina . . . .	8
2.1. Clasificación de los diferentes trabajos de detección de estructuras retinianas.	21
2.2. Clasificación de los diferentes trabajos de segmentación de estructuras retinianas. . . . .	40
3.1. Diagrama de bloques de la metodología utilizada en detección. . . . .	61
3.2. <i>Templates</i> de disco óptico (DO) utilizados para cada canal RGB . . . . .	65
3.3. <i>Templates</i> de mácula utilizados para cada canal RGB . . . . .	66
3.4. Ejemplo de detección de la red de vasos principales . . . . .	69
3.5. Diagrama de bloques para obtención de blob de vasos . . . . .	70
3.6. <i>Templates</i> de haz vascular (HV) utilizados para cada tipo de ojo . . . . .	72
3.7. Diagrama de bloques para obtención de los diferentes blobs candidatos . . . .	73
3.8. Diagrama de bloques asociado a la metodología usada para la segmentación simultánea de las estructuras #1 y #2. . . . .	79
3.9. Ejemplo que muestra el etiquetado de los puntos que pertenecen al contorno de DO . . . . .	84
3.10. Modelo de DO ( $\mathbf{M}_{DO}$ ) . . . . .	85
3.11. Modelo de fóvea ( $\mathbf{M}_F$ ) . . . . .	90
3.12. Representación esquemática del modelo relacional DO-fóvea ( $\mathbf{M}_{DO-F}$ ). . . . .	93
3.13. Acoplamiento del método de detección al método de segmentación . . . . .	99
4.1. Dos ejemplos con resultados intermedios y final del proceso de detección (I) . .	110

4.2. Dos ejemplos con resultados intermedios y final del proceso de detección (II)	. 111
4.3. Detección correcta del DO tras considerar cuasi-equivalentes las posiciones de DO y HV.	. . . . . 123
4.4. Ejemplos de segmentación obtenidos para diferentes imágenes MESSIDOR.	. . 132

# Índice de tablas

2.1. Categorización de diferentes trabajos de detección atendiendo a varios descriptores . . . . .	37
2.1. Categorización de diferentes trabajos de detección atendiendo a varios descriptores (Cont.) . . . . .	38
2.1. Categorización de diferentes trabajos de detección atendiendo a varios descriptores (Cont.) . . . . .	39
2.2. Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores . . . . .	55
2.2. Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores (Cont.) . . . . .	56
2.2. Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores (Cont.) . . . . .	57
2.2. Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores (Cont.) . . . . .	58
3.1. Área retinal para diferentes valores de ángulos del campo de visión (FOV) . . . . .	63
3.2. Descripción de cada uno de los genes codificados en un cromosoma. . . . .	94
3.3. Cálculo del intervalo de restricción ( $X_o, X_f$ ) con el 95 % de valores dentro del intervalo . . . . .	96
4.1. Comparativa de los resultados de detección de DO y mácula obtenido en MES-SIDOR . . . . .	107

4.2. Comparativa de los resultados de detección de DO en MESSIDOR atendiendo a diferentes grados de patologías . . . . .	108
4.3. Comparativa de los resultados de detección de mácula en MESSIDOR atendiendo a diferentes grados de patologías . . . . .	108
4.4. Comparativa de los resultados de detección de DO, mácula y HV para otras bases de imágenes públicas . . . . .	112
4.5. Tasa de éxito en la base de imágenes MESSIDOR para cada estructura anatómica, según la cantidad de CRinter-e inyectado en el proceso de búsqueda . . .	119
4.6. Comparativa con otros métodos considerando hardware, lenguaje de programación, tiempo de cálculo, número y tipos de estructuras detectadas. . . . .	121
4.7. Valor de parámetros utilizados en el algoritmo DE . . . . .	127
4.8. Influencia de usar o no el conocimiento relacional inter-estructura (CRinter-e) en el proceso de segmentación. . . . .	128
4.9. Comparativa de realizar segmentación de estructuras de forma simultánea o secuencial. . . . .	130
4.10. Comparativa con otros métodos de segmentación simultánea de DO y fóvea. .	130
4.11. Comparativa con otros métodos competitivos de la literatura que solo abordan el problema de segmentación de DO. . . . .	131
4.12. Comparativa (en tanto por ciento) con otros métodos competitivos de la literatura que solo abordan el problema de segmentación de fóvea. . . . .	131
4.13. Comparativa con otros métodos competitivos que solo abordan la segmentación de DO (base de imágenes ONHSD). . . . .	134
4.14. Resultados de segmentación de fóvea para ONHSD. . . . .	134
4.15. Comparativa con otros métodos de segmentación simultánea de DO y fóvea, considerando el hardware, lenguaje de programación y tiempo de computación. .	138
4.16. Influencia de usar o no los resultados del método de detección en el proceso de segmentación (MESSIDOR). . . . .	139
4.17. Influencia de usar o no los resultados del método de detección en el proceso de segmentación (ONHSD). . . . .	140

# Lista de acrónimos

Esta es una lista de acrónimos utilizados en esta tesis:

**ROI** Region Of Interest

**ACO** Ant Colony Optimization

**AE** Algoritmo Evolutivo

**AGF** Angiografía Fluorescénica

**AHE** Adaptive Histogram Equalization

**ASM** Active Shape Model

**CCD** Charge-Coupled Device

**CIF** Convergency Index Filter

**CMA-ES** Covariance Matrix Adaptation Evolution Strategy

**CRinter-e** Conocimiento Relacional inter-estructura

**CRintra-e** Conocimiento Relacional intra-estructura

**DE** Differential Evolution

**DMF** Directional Matched Filtering

**DO** Disco Óptico

**FOV** Field Of View

**FPF** False Positive Fraction

**FRS** Fast Radial Symmetry

**GMP** Generalized Motion Pattern

**GPU** Graphics Processing Unit

**GRD** Grado de Retinopatía Diabética

**HRT** Heidelberg Retina Tomography

**HSI** Hue, Saturation and Intensity (espacio de color)

**HV** Haz Vascular

**K-NN** k-Nearest Neighbors Algorithm

**LSM** Level Set Method

**MRF** Markov Random Field

**OCT** Optical Coherence Tomography

**PCA** Principal Component Analysis

**PH** Puntos Hipótesis

**RANSAC** Random Sample Consensus

**RDNP** Retinografía Diabética No Proliferativa

**RDP** Retinografía Diabética Proliferativa

**REM** Riesgo de Edema Macular

**RGB** Red, Green and Blue (espacio de color)

**ROI** Region Of Interest

**SBF** Sliding Band Filter

**SEF** Super-Elliptical Filter

**SVM** Support Vector Machines

**TAN** Topological Active Network

**TM** Template Matching

**TPF** True Positive Fraction



# Capítulo 1

## Introducción

Dentro de este capítulo, la motivación de la tesis está expuesta en la sección 1.1. Luego, en la sección 1.2, se describe el alcance de tesis, incluyendo: la descripción del problema abordado, las limitaciones de la tesis, así como las hipótesis y objetivos de investigación. A continuación, en la sección 1.3, se describe el material utilizado durante el desarrollo de la tesis. Finalmente, en la sección 1.4, se presenta la estructura de esta tesis.

### 1.1. Motivación

La fotografía de fondo de ojo permite tomar imágenes detalladas de la retina. Esto permite registrar el estado visual de diferentes estructuras retinianas, proporcionando una valiosa información para el diagnóstico de enfermedades oculares. La gran calidad de estas imágenes, el realce de contraste o el uso de filtros de imagen, permiten destacar elementos que a simple vista son difíciles de detectar mediante un oftalmoscopio (dispositivo que permite la visión directa del fondo de ojo). Además, el registro temporal de este tipo de imágenes permite realizar el seguimiento de ciertas enfermedades oculares. Por tanto, el análisis automático de las imágenes de retina mediante un sistema de visión puede constituir una estimable ayuda para el diagnóstico.

Son varias las enfermedades oculares degenerativas que no muestran claros signos o síntomas en sus primeras etapas, pero que, por las graves consecuencias en el paciente que las

sufre, resulta de vital importancia una detección precoz de las mismas. Así, la *retinopatía diabética* viene producida por una complicación en enfermos con diabetes, y constituye la principal causa de ceguera irreversible en la población activa del mundo desarrollado [Taylor and Keeffe, 2001]. En la etapa más temprana, denominada *retinopatía diabética no proliferativa* (RDNP), aparecen distintos signos patológicos visuales: microaneurismas, microhemorragias y exudados duros entre otros. En la etapa final, denominada *retinopatía diabética proliferativa* (RDP), la mayoría de los vasos se encuentran taponados, por lo que el flujo sanguíneo se ralentiza, facilitando la proliferación de nuevos vasos en la zona del disco óptico (neovascularización). De otro lado, el *glaucoma* es la segunda causa de ceguera irreversible a nivel mundial. En 2013, más de 64 millones de personas de todo el mundo en la franja de edad 40-80 sufrieron esta enfermedad, con una estimación de alcanzar los 111 millones en 2020 [Maheshwari et al., 2017]. Normalmente es causada por una presión ocular elevada (hipertensión) debido a la dificultad de drenaje del fluido intraocular (dificultad generada por causas hereditarias o incluso como consecuencia de la retinopatía diabética). Uno de los signos característicos de este tipo de patología es el aumento de la excavación (zona central del disco óptico). Adicionalmente, el denominado *edema macular* es una inflamación de la zona macular que, si no es tratado, puede causar pérdidas de la visión e incluso ceguera. Los signos visuales característicos de este tipo de patología corresponde a la existencia de exudados duros en la zona macular. Finalmente, la denominada *degeneración macular asociada a la edad* afecta a la zona de la retina responsable de la visión detallada (visión central), esto es, a la mácula. No afecta a la visión periférica por lo que normalmente no causa ceguera total. Los signos visuales asociados a esta enfermedad son la proliferación de drusas bajo la retina cerca del área macular.

Durante el examen ocular, los oftalmólogos buscan regiones específicas y determinados patrones para identificar posibles marcas de enfermedad. En este proceso, la identificación y análisis del estado de las principales estructuras anatómicas retinianas es clave. Por lo tanto, si se quiere implementar un sistema de ayuda al diagnóstico de patologías oculares basado en imágenes de retina, será necesario, como paso previo, abordar la detección y segmentación automática de las estructuras anatómicas más relevantes.

De un lado, existen muchos argumentos en la literatura que justifican la importancia de la detección de estructuras retinianas. Así, los resultados de detección de las diferentes estructuras anatómicas pueden utilizarse como puntos de referencia (*landmarks*) para otros sistemas de diagnóstico automatizados con el fin de detectar signos de enfermedades [Mahfouz and Fahmy, 2010, Gegundez et al., 2013, Giachetti et al., 2013, Mittal and Sivaswamy, 2015, Foracchia et al., 2004, Yu et al., 2012]. Por ejemplo, la detección del disco óptico es un requisito previo para el cálculo de algunos índices diagnósticos importantes para la denominada *retinopatía hipertensiva/esclerótica* [Foracchia et al., 2004, Hubbard et al., 1999]. De la misma manera, la presencia de una lesión brillante/oscura próxima a la mácula indica una mayor probabilidad de alteración de la visión y, por tanto, la necesidad de una atención médica inmediata [Mittal and Sivaswamy, 2015]. Adicionalmente, la búsqueda de patrones visuales de patología en una imagen retiniana puede suponer la necesidad de aplicar una máscara a determinadas estructuras anatómicas para evitar confundirlas con dichos patrones [Niemmeijer et al., 2007]. Así, por ejemplo, dado que el disco óptico puede confundirse fácilmente con lesiones grandes de tipo exudativo, por presentar un color similar, la detección correcta del primero es importante para excluirlo del conjunto de posibles lesiones [Foracchia et al., 2004]. Un ejemplo de esto lo encontramos en [Osareh, 2004], donde, en una primera etapa, se detecta y elimina el disco óptico con el objeto de disminuir la tasa de falsos positivos en la búsqueda y clasificación de exudados. Un argumento similar se aplica a la fovea cuando el objetivo es detectar automáticamente hemorragias retinianas [Medhi and Dandapat, 2016]. Finalmente, la detección de una estructura retiniana puede ser un paso adicional al proceso de segmentar dicha estructura, no solo por el hecho de reducir el tiempo de computación, sino porque puede favorecer la precisión y especificidad del proceso de segmentación [Cao et al., 2013]. Por ejemplo, existen muchos enfoques en los que la detección de disco óptico se utiliza como paso previo a su segmentación [Lowell et al., 2004, Carmona et al., 2008, Aquino et al., 2010, Welfer et al., 2010, Marin et al., 2015]. Por último, la detección también se puede utilizar para distinguir automáticamente el tipo de ojo (izquierdo y derecho) mostrado en una imagen de retina [Mahfouz and Fahmy, 2010].

De otro lado, la segmentación de estructuras anatómicas retinianas es, si cabe, aún más

importante que la detección, ya que constituye una identificación más precisa de cada estructura. Su importancia radica en el hecho de que, para diagnosticar ciertas enfermedades, no basta solo con conocer la localización aproximada de las estructuras involucradas, sino que es necesario conocer la forma y límites de la misma. Además, los cambios en el tiempo de la forma o límites pueden dar información sobre la evolución de ciertas enfermedades. En la literatura también se muestran ejemplos que resaltan la importancia de la segmentación de estructuras retinianas. Así, una segmentación detallada del disco óptico constituye un paso previo fundamental para un sistema automático de ayuda al diagnóstico [Díaz-Pernil et al., 2016]. Más específicamente, en diagnóstico de glaucoma, es importante conocer la forma, color y área del disco óptico [Aquino et al., 2010], o incluso la relación de área entre este último y la excavación [Carmona et al., 2008, Tang et al., 2003, Cheng et al., 2011, Morales et al., 2013, Haleem et al., 2013, Bharkad, 2017, Lu, 2011, Noor et al., 2013]. Igualmente, la segmentación del disco óptico es esencial para detectar anomalías asociadas a esta estructura, como puede ser la neovascularización del disco óptico o el papiledema [Yu et al., 2012]. La localización precisa del límite del disco óptico también es muy útil para el diagnóstico de la RDP, debido a la presencia de neovascularización en la vecindad de esta estructura [Kande et al., 2008]. Igualmente, la segmentación del disco óptico es una etapa importante para el cálculo de índices relacionados con los cambios vasculares, como la denominada *relación arteriolar-venular* [Dashtbozorg et al., 2015]. De otro lado, la distancia entre ciertas lesiones (como los exudados duros) y la fovea son de importancia clínica para medir el daño producido y por tanto clasificar el grado de edema macular [Kao et al., 2014, Medhi and Dandapat, 2016, Gegundez et al., 2013, Welfer et al., 2011, Aquino, 2014], siendo el daño más severo cuanto más cerca se encuentren las lesiones del centro de la mácula. La segmentación de los vasos sanguíneos es importante para la detección de diferentes signos asociados a la RDNP, tal como la neovascularización [Ciulla and Amador, 2003, Kauppi et al., 2007] o los rosarios venosos (*venous beading*) [Vallabha et al., 2004], o como indicador de pacientes con hipertensión [Wong et al., 2004, Ruggeri et al., 2007] ya que el estrechamiento y/o variación de la tortuosidad vascular es un claro síntoma de retinopatía hipertensiva. Por último, también relacionado con el glaucoma, está la tarea de seguimiento y evolución temporal de los

vasos retinianos a fin de encontrar microaneurismas. En este sentido, se utiliza la posición de los vasos dentro del área del disco óptico como semillas para el seguimiento de los mismos [Siddalingaswamy and Prabhu, 2010, Roychowdhury et al., 2016].

A pesar de la existencia en la literatura de una gran cantidad de referencias dedicadas a la detección y segmentación de estructuras retinianas, ambos problemas siguen siendo hoy día un problema abierto, donde los esfuerzos van dirigidos principalmente hacia la implementación de métodos cada vez más rápidos y robustos.

## 1.2. Alcance de la tesis

En esta sección se presenta el alcance de la tesis, describiendo en primer lugar el problema abarcado. A continuación se exponen las limitaciones de la tesis, y finalmente, las hipótesis y objetivos de investigación propuestos.

### 1.2.1. Descripción del problema

Además de la observación directa de la retina mediante un oftalmoscopio, esta puede ser examinada indirectamente a partir de una fotografía digital de fondo de ojo. Existen diferentes técnicas para capturar la retina aprovechando las propiedades reflectantes del fondo de ojo y empleando diferentes tipos de luz y dispositivos de captura. La más directa y económica es la que emplea el retinógrafo. Básicamente, el retinógrafo es una cámara digital basada en CCD (del inglés *Charge-Coupled Device*) que produce imágenes digitales en formato, normalmente, RGB. No obstante existen otras técnicas basadas en un sistema de barrido con haz de láser de infrarrojo como, por ejemplo, la denominada tomografía de coherencia óptica (OCT, del inglés *Optical Coherence Tomography*) o la tomografía de retina de Heidelberg (HRT, del inglés *Heidelberg Retina Tomography*). Todos los métodos mencionados constituyen una manera no invasiva e inocua de examinar las estructuras anatómicas oculares. Por el contrario, existe otras técnicas como, por ejemplo, la denominada *angiografía fluoresceínica* (AGF), que sí es invasiva, ya que requiere inyectar al paciente un contraste para poder capturar las imágenes.

En esta tesis se plantea el problema de la detección y segmentación de las principales estructuras anatómicas en imágenes digitales de retina (ver Fig. 1.1), es decir, disco óptico, red de vasos, haz vascular y mácula/fóvea. A la hora de identificar una estructura anatómica en una imagen de retina, se debe distinguir entre detección y segmentación. De un lado, la tarea de detección de estructuras anatómicas se define como la búsqueda del punto que representa el centro del área abarcada por la estructura o, en su defecto, de un punto que está contenido en el área que abarca dicha estructura. Por otro lado, la tarea de segmentación implica la búsqueda de todos los puntos que forman el área o perímetro de la estructura. En este punto, hay que decir que no existe una distinción clara en la literatura entre detección y segmentación de mácula/fóvea. Esto es debido a que la frontera de ambas estructuras es mucho más difusa que la que tienen los vasos o el disco óptico. Sin embargo, se puede hablar indistintamente de centro de mácula o de fóvea porque es un punto común en ambas estructuras. Así, tanto los trabajos de detección como segmentación de mácula/fóvea suelen utilizar como criterio de acierto la distancia existente desde el centro detectado/segmentado hasta el centro real de ambas estructuras. En esta tesis, a fin de establecer una frontera clara entre detección o segmentación de mácula/fóvea, se establecerá un umbral igual a  $\frac{1}{2}R_{DO}$  (siendo  $R_{DO}$  el radio medio del disco óptico), de tal manera que un método será catalogado de detección o segmentación dependiendo de si el criterio utilizado para medir la distancia entre el centro obtenido y el real es, respectivamente, mayor o menor-igual que el umbral mencionado. Además, para mantener la coherencia y claridad en las explicaciones, y dado que la mácula es una estructura más amplia que incluye a la fóvea, nos referiremos a la primera cuando hablemos de detección y, a la segunda, cuando hablemos de segmentación.

Aunque ambas tareas son importantes por separado, una alta tasa de éxito en el proceso de detección facilitará enormemente el éxito de la tarea de segmentación, ya sea de la misma estructura o de otra estructura que utiliza como referencia la detección de la primera. Las etapas siguientes implicadas en la construcción de un sistema automático de ayuda al diagnóstico basado en imágenes también serían favorecidas por tales éxitos.

El *disco óptico* (DO), también llamado *papila* o *cabeza del nervio óptico*, tiene forma ovalada y está localizado en el área donde convergen las fibras nerviosas retinianas. Dicha

área constituye el comienzo del nervio óptico, encargado de transmitir la información visual desde la retina hasta el cerebro. En el centro del DO está la *excavación*, que es la depresión que dejan las fibras nerviosas cuando cambian la trayectoria para dirigirse a la porción intracraneal del nervio óptico. En ausencia de patologías y ruido, la excavación aparece como la parte más brillante de la imagen. La zona que rodea la excavación y alcanza el borde externo del DO se llama *anillo neuroretiniano*. En condiciones normales, este anillo tiene una coloración anaranjada o rosácea y su perímetro externo delimita el contorno del DO. Esta estructura puede dividirse geoméricamente en cuatro cuadrantes comprendidos en los intervalos  $[45^\circ, 135^\circ]$  (zona temporal),  $[135^\circ, 225^\circ]$  (zona inferior),  $[225^\circ, 315^\circ]$  (zona nasal) y  $[315^\circ, 45^\circ]$  (zona superior) para ojo izquierdo. En el caso de ojo derecho, se intercambian las zonas temporal y nasal. Cualquier cambio en la forma, profundidad o color del DO, o en la relación de diámetros DO-excavación, es utilizado por los oftalmólogos para diagnosticar diferentes patologías retinianas. De aquí que la detección y segmentación del DO sea de crítica importancia en un sistema automático de apoyo al diagnóstico de enfermedades oculares.

La *red de vasos* está formada por arterias y venas que, emergiendo desde la zona central del DO, se despliegan y ramifican por toda la retina siguiendo un patrón en forma de árbol. Existen diferencias morfológicas entre los vasos principales, los cuales son más anchos y lineales, y los vasos más finos (capilares), los cuales son más estrechos y curvos. El árbol vascular puede considerarse cuasi-simétrico en relación al eje que une los centros de DO y fovea, y que se corresponde aproximadamente con el denominado *rafe horizontal*, delimitando así las denominadas arcadas superior e inferior que corresponden, respectivamente, a la parte superior e inferior de la red de vasos.

El *haz de vasos* o *haz vascular* (HV) está formado por los vasos que se encuentran dentro del perímetro del DO. El punto que une las arcadas superior e inferior, situado dentro del DO, constituye el centro del HV. Dado que dicho centro se encuentra normalmente muy cerca del centro del DO, tal y como se verá en el siguiente capítulo, existen trabajos en la literatura que utilizan la detección del primero como detección aproximada del segundo o viceversa.

La *mácula* tiene forma redondeada y se encuentra delimitada por la zona temporal del DO y las arcadas superior e inferior de la red vascular. De otro lado, la *fovea* es una pequeña

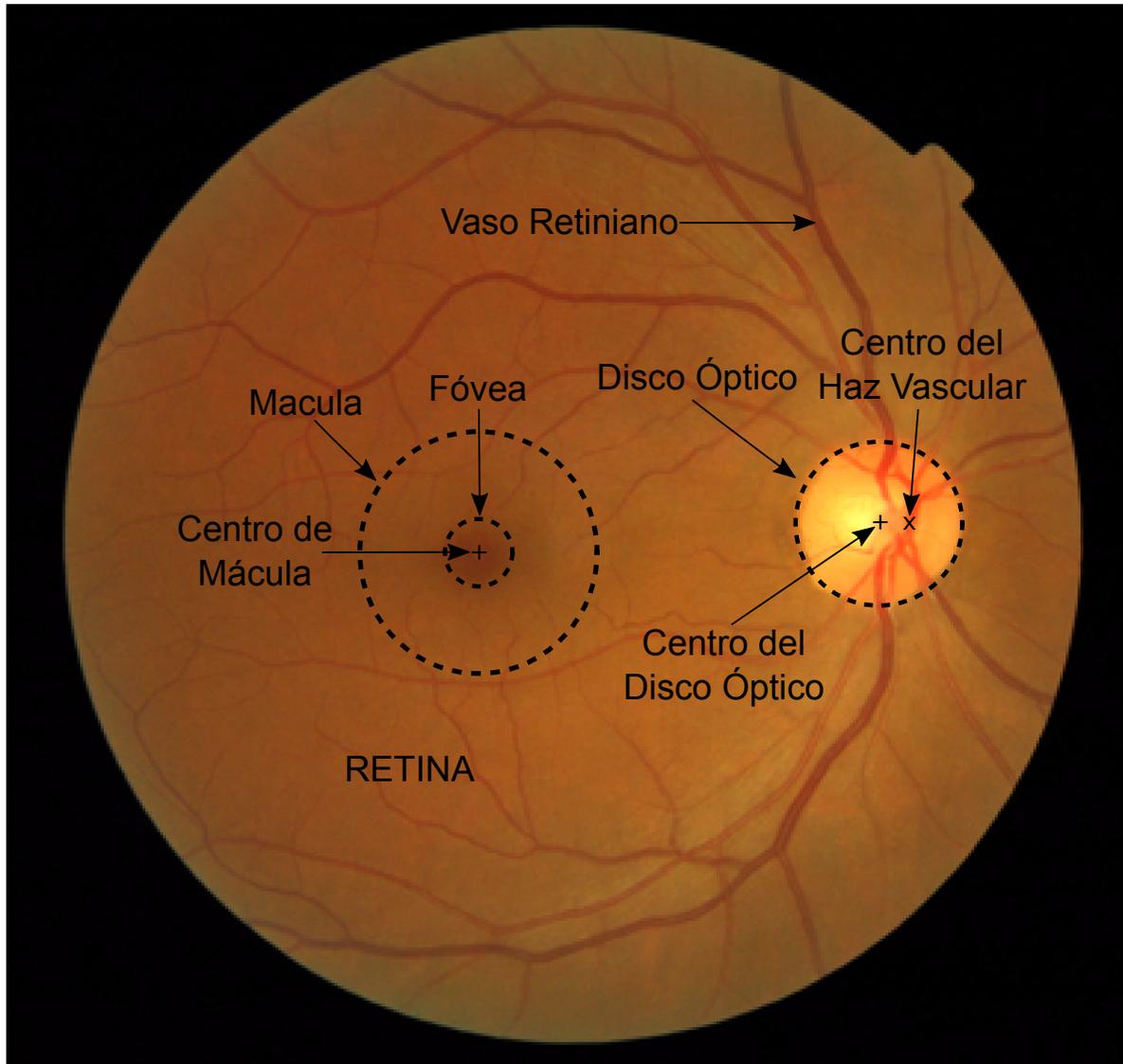


Figura 1.1: Principales estructuras anatómicas a considerar en una imagen de retina (ojo derecho).

depresión avascular situada en la parte central de la mácula. Es la zona especializada en la visión fina de los detalles (o visión central) y nos sirve entre otras cosas para poder leer y distinguir las caras de las personas. En ausencia de patologías oculares y de ruido, la fovea se muestra en una imagen de retina como la zona más oscura. Por la importancia de su función, la detección y segmentación de la fovea (mácula), al igual que ocurre con el DO, juega un papel clave en cualquier sistema de ayuda al diagnóstico.

La detección y segmentación de las estructuras anatómicas se puede ver dificultada por la presencia de patrones visuales asociados a diferentes patologías retinianas. Así, las denominadas *lesiones oscuras* (también llamadas *lesiones rojas*) comprenden los *microaneurismas* y las *hemorragias*. Los microaneurismas son pequeñas inflamaciones vasculares que acumulan sangre. Aparecen en zonas cercanas a los vasos, como un pequeño punto apenas perceptible de color rojo vivo y de bordes netos. Cuando los microaneurismas se rompen, dejan escapar pequeñas cantidades de sangre, produciéndose así las hemorragias. Estas últimas tienen un aspecto muy parecido al de los microaneurismas, pero pueden ser mucho más grandes. La presencia de las lesiones oscuras en la imagen pueden dificultar en gran medida la detección o segmentación de la mácula o de la fovea, respectivamente. De otro lado, también existe lo que se denominan *lesiones brillantes*, que comprenden los *exudados duros*, los *exudados algodinosos* y las *drusas*. Los exudados duros se producen por rotura de inflamaciones vasculares que acumulan grasa (producto de desecho del metabolismo). Estas estructuras presentan un aspecto amarillento, de bordes irregulares pero precisos. De otra parte, los exudados algodinosos, como su propio nombre indica, tienen el aspecto de manchas blancas y redondeadas, pero con límites no definidos, siendo más grandes que los exudados duros. Este tipo de exudados se generan a consecuencia de la falta de oxígeno en las fibras nerviosas retinales, las cuales se deterioran acumulando agua (de ahí el aspecto blanquecino). Las drusas aparecen durante el proceso de envejecimiento, son pequeños depósitos de grasa que se acumulan bajo la retina, tienen un aspecto redondeado y color amarillento. Debido a su color, pueden confundirse con exudados duros pero, a diferencia de estos, su coloración amarilla es tenue y apagada. Existe otro grupo de lesiones brillantes, denominadas *atrofias peripapilares* y *estafilomas posteriores*, que están normalmente localizadas en la vecindad del DO. La atrofia peripapilar se produce

por una alteración en los vasos y fibras nerviosas cercanos al DO. Esta lesión provoca decoloración del DO y cambios en su estructura, normalmente caracterizada por un aumento del perímetro del mismo. El estafiloma posterior consiste en una deformación de la pared ocular posterior, debido a un proceso de adelgazamiento y expansión de la esclerótica. Se manifiesta como una gran zona brillante de bordes redondeados y precisos que rodea al DO, pudiendo llegar hasta la mácula. La presencia de lesiones brillantes en la imagen suelen dificultar tanto la detección como la segmentación del DO.

La detección o segmentación de cualquier estructura anatómica ocular es una tarea compleja. Las imágenes contienen frecuentemente ruido introducido en el proceso de adquisición (presencia de artefactos visuales, insuficiente contraste de la imagen o una inadecuada iluminación). Igualmente, la existencia de los diferentes patrones visuales de patologías retinianas descritos anteriormente pueden dificultar o confundir la identificación de las estructuras anatómicas a detectar o segmentar. Si lo anterior no fuese suficiente, la variabilidad anatómica de cada individuo dificulta aún más el proceso.

### **1.2.2. Limitaciones**

Para delimitar el alcance de esta tesis es preciso decir que aquí solo se abordará el problema de la detección y segmentación de las principales estructuras anatómicas retinianas, quedando fuera de dicho alcance la detección de patrones visuales asociados a diferentes patologías oculares. En cuanto al tipo de imágenes, el estudio se restringe a imágenes digitales obtenidas con retinógrafo y formato RGB. Queda también fuera del propósito de esta tesis el estudio e implementación de un sistema automático de ayuda al diagnóstico de patologías oculares.

### **1.2.3. Hipótesis y objetivos de investigación**

Una de las principales carencias observadas en aquellos trabajos que abordan la detección (segmentación) de una única estructura anatómica retiniana es que no se maximiza el uso del conocimiento relacional existente con otras estructuras, es decir, solo se utiliza el conocimiento de la propia estructura durante el proceso de detección (segmentación) o, si se usa

el conocimiento relacional con otras estructuras, este es limitado. Existen también trabajos en los que la detección (segmentación) de una estructura se basa en la detección (segmentación) de otra estructura previamente detectada (segmentada). En estos casos, aunque sí se utiliza el conocimiento relacional existente entre ambas estructuras, un fallo en la detección (segmentación) de la estructura usada como referencia, producirá irremediablemente un fallo en la detección (segmentación) de la segunda. En ambos casos, la robustez del proceso de detección (segmentación) se ve afectado.

De aquí en adelante, al conocimiento relacionado con la propia estructura se le denominará *conocimiento relacional intra-estructura* (CRintra-e) y, al conocimiento relacional existente entre dos o más estructuras, *conocimiento relacional inter-estructura* (CRinter-e). En esta tesis se plantea no solo el uso de ambos tipos de conocimiento, CRintra e inter-e, sino también que su uso se haga de forma conjunta con el objeto de realizar la detección (segmentación) simultánea de dos o más estructuras retinianas. Este planteamiento da lugar a las siguientes hipótesis de investigación:

1. El uso del CRintra-e favorece la detección (segmentación) de verdaderos positivos, pero no evita la presencia de falsos positivos (sobre todo en aquellas imágenes que contengan ruido o signos visuales de patologías). Por tanto, el uso adicional del CRinter-e favorecerá la eliminación de los falsos positivos y, consecuentemente, aumentará las probabilidades de éxito durante el proceso de detección (segmentación).
2. La detección (segmentación) simultánea de  $n$  estructuras a partir del uso conjunto del CRintra e inter-e favorecerá la elección de aquella  $n$ -tupla de estructuras que mejor explique ambos tipos de conocimiento y, por tanto, también aumentará la bondad del proceso de detección (segmentación).
3. La incorporación adicional del conocimiento de detección simultánea (salida del proceso de detección) al proceso de segmentación favorecerá la bondad de este último.

En el contexto del problema definido y de las hipótesis planteadas, se establecen los siguientes objetivos a conseguir en el desarrollo de esta tesis:

1. Crear un modelo relacional para cada estructura anatómica a partir de su respectivo CR<sub>intra-e</sub> asociado. De aquí en adelante, este tipo de modelo se denominará *modelo relacional intra-e* o, simplemente, *modelo intra-e*.
2. Crear un modelo relacional del conjunto de estructuras a detectar (segmentar) que incluya gran parte del CR<sub>inter-e</sub> existente entre todas ellas. De aquí en adelante, este tipo de modelo se denominará *modelo relacional inter-e* o, simplemente, *modelo inter-e*. Dado que la tarea de detección es menos exigente que la de segmentación, los *modelos intra e inter-e* exigidos para la primera serán menos complejos que los exigidos para la segunda.
3. Establecer una metodología de detección simultánea que establezca cómo hacer un uso conjunto de la colección de modelos relacionales definidos en los dos primeros objetivos. La metodología debe ser independiente del tamaño de la imagen de entrada.
4. Establecer una metodología de segmentación simultánea que establezca cómo hacer un uso conjunto de la colección de modelos relacionales definidos en los dos primeros objetivos. Ambas metodologías, tanto la de detección como la de segmentación, deben ser independientes del tamaño de la imagen de entrada.
5. Utilizando la metodología descrita en el punto 3, implementar y evaluar un método que permita la detección simultánea de las principales estructuras anatómicas retinianas (DO, mácula, HV y red de vasos) a partir del uso conjunto de los modelos intra e inter-e.
6. Utilizando la metodología descrita en el punto 4, implementar y evaluar un método que permita la segmentación simultánea de las principales estructuras anatómicas retinianas (DO y fovea) a partir del uso conjunto de los modelos intra e inter-e.
7. Implementar una estrategia que permita al método de segmentación aprovechar los resultados obtenidos por el método de detección, de tal forma que estos últimos se usen como una entrada adicional en el método de segmentación.

### 1.3. Materiales

A lo largo de esta tesis, se han utilizado cinco bases de imágenes públicas a la hora de realizar los diferentes experimentos de detección y segmentación: MESSIDOR [MESSIDOR, 2012, Decenciere et al., 2014], DRIONS-DB [Carmona et al., 2008, DRIONS-DB, 2013], ONHSD [ONHSD, 2013, Lowell et al., 2004], DIARETDB1 [DIARETDB1, 2015, Kauppi et al., 2007] y DRIVE [DRIVE, 2004, Staal et al., 2004]. La base de imágenes MESSIDOR contiene 1200 imágenes de retina, formato RGB, 8 bits/píxel, FOV = 45° (FOV, del inglés *field of view*) y tres tamaños distintos: 1440 × 960, 2240 × 1488 y 2304 × 1536 píxeles. Las anotaciones de *ground truth* para el contorno del DO en esta base de imágenes están disponibles en la web [MESSIDOR, 2012]. Las anotaciones<sup>1</sup> de los centros de fovea para 1136 imágenes de MESSIDOR fueron creadas y publicadas por la Universidad de Huelva. Las restantes 64 anotaciones, hasta las 1200, fueron amablemente proporcionadas por los autores de [Gegundez et al., 2013]. La base de imágenes DRIONS consta de 110 imágenes de retina, formato RGB, 8 bits/píxel, FOV = 40° y un tamaño de 600 × 400 píxeles. En la web de esta base de imágenes [DRIONS-DB, 2013] solo se proporciona el *ground truth* del contorno de DO. Específicamente, cada centro de DO utilizado como referencia en la tarea de detección se obtuvo como resultado del promedio de los centros geométricos de dos contornos de DO trazados por dos expertos diferentes. La base de imágenes ONHSD contiene 99 imágenes de retina, formato RGB, 8 bits/píxel, FOV = 45° y tamaño de 640 × 480 píxeles. Al igual que ocurre con DRIONS, en la web de esta base de imágenes [ONHSD, 2013] solo se proporciona el *ground truth* del contorno de DO. En concreto, cada centro de DO utilizado como referencia en la tarea de detección se obtuvo como resultado del promedio de los centros geométricos de cuatro contornos, cada uno de ellos trazado por un experto diferente, y por otra parte, los centros de fovea en esta base de imágenes fueron anotados por un especialista en oftalmología adscrito al Hospital Clínico San Carlos de Madrid. La base de imágenes DIARETDB1 contiene 89 imágenes de retina, formato RGB, FOV = 50° y tamaño de 1500 × 1152 píxeles. En la web de esta base de imágenes [DIARETDB1, 2015], no hay *ground truth* ni para

---

<sup>1</sup>[http://www.uhu.es/retinopathy/muestras/Provided\\_Information.zip](http://www.uhu.es/retinopathy/muestras/Provided_Information.zip)

la fovea ni para el DO. Aquí, los centros de DO fueron anotados por el mismo especialista en oftalmología mencionado anteriormente y las anotaciones de fovea fueron obtenidas de [Aquino, 2014]. Finalmente, la base de imágenes DRIVE contiene 40 imágenes de retina, formato RGB, 8 bits/píxel, FOV = 45° y un tamaño de 768 × 584 píxeles. Esta última base de imágenes solo se utilizó para abordar la tarea de detección de la red de vasos, ya que, desde su página web, solo está públicamente disponible el ground truth de dicha estructura. Todas las bases de imágenes mencionadas contienen tanto ojos sanos como ojos con diferentes tipos de retinopatías. Se puede consultar una descripción más detallada de cada base de imágenes en sus respectivas referencias.

### 1.4. Estructura de la tesis

A continuación presentamos un breve resumen del contenido de cada uno de los capítulos y anexos de los que se compone esta tesis:

En el **capítulo 1** (capítulo actual) se expone la motivación y alcance de esta tesis, así como los materiales utilizados.

Seguidamente, en el **capítulo 2**, se hace una revisión de los trabajos existentes en la literatura sobre detección y segmentación de estructuras anatómicas retinianas. También se muestra una taxonomía de estos trabajos basada principalmente en el uso conjunto o no del CRintra e inter-e y en el carácter aislado, secuencial o simultáneo del proceso de detección/segmentación.

A continuación, el **capítulo 3** presenta y describe las metodologías de detección y segmentación propuestas en esta tesis. Igualmente se describen los pasos seguidos para: (i) la construcción de los diferentes modelos intra-e e inter-e asociados a las diferentes estructuras retinianas, y (ii) la implementación de los métodos de detección y segmentación basados en sendas metodologías propuestas.

En el **capítulo 4**, se exponen los experimentos realizados para evaluar diferentes aspectos relacionados tanto con el método de detección como con el de segmentación, utilizando para ello diferentes bases de imágenes públicas. Se muestran los resultados obtenidos y una

detallada discusión de los mismos.

Finalmente, en el **capítulo 5**, se enumeran las conclusiones derivadas de esta tesis, así como las posibles líneas futuras de investigación.

Adicionalmente, el **Anexo A** proporciona información de las diferentes publicaciones obtenidas como resultado del trabajo realizado en esta tesis y el **Anexo B** muestra las conclusiones y líneas de trabajo futuro redactadas en inglés.



# Capítulo 2

## Estado del arte

En este capítulo se describen y analizan diferentes trabajos existentes en la literatura que están dedicados a abordar las dos tareas principales que son objeto de esta tesis: la detección y segmentación de estructuras retinianas. Previamente, también se realiza una descripción de diferentes técnicas de preprocesado de imagen que han sido típicamente usadas en el contexto que nos ocupa y que han estado dirigidas a facilitar la implementación de métodos que abordan las dos tareas mencionadas.

### 2.1. Preprocesado en imágenes de retina

En muchos trabajos de detección (segmentación) de estructuras anatómicas retinianas, se realiza una etapa de preprocesamiento de la imagen de entrada. Esta etapa pretende fundamentalmente corregir aspectos de la imagen y/o eliminar posible ruido que, en ambos casos, pudiese dificultar las etapas de detección (segmentación). Además, otros métodos de preprocesamiento van encaminados a reducir el coste computacional de dichas etapas. Estas técnicas podrían clasificarse en distintos grupos: mejora de contraste, actuación sobre el espacio de color, reducción de ruido, reescalado de imagen y reducción del área de trabajo.

Muchas imágenes de fondo de ojo presentan un mal contraste que dificulta el proceso de detección (segmentación). Por lo tanto, se hace necesario aplicar técnicas de realce de contraste, que actúan sobre la imagen completa o en una zona (o estructura) concreta de la misma.

Los principales métodos de realce son la ecualización del histograma y la corrección de luminosidad no uniforme. Así, en algunos trabajos se mejora el contraste mediante ecualización del histograma [Petsatodis et al., 2004, Girard et al., 2016]. Una mejora del método anterior es el uso de la ecualización del histograma adaptativo (AHE, del inglés *Adaptive Histogram Equalization*), donde para modificar la intensidad de cada píxel se considera una ventana local a dicho píxel (en vez de usar valores de píxel como se hace en la técnica no adaptativa). En [Sinthanayothin et al., 1999, Youssif et al., 2008, Zhu et al., 2010, Salazar-Gonzalez et al., 2014] se aplica esta técnica de ecualización adaptativa. La corrección de luminosidad también se aplica en [Hoover and Goldbaum, 2003, Youssif et al., 2008]. En otros trabajos, en vez de realizar un realce de contraste a la imagen completa, este se aplica a fin de resaltar una estructura concreta como, por ejemplo, en [Fleming et al., 2007], donde se realza la red de vasos como paso previo a su detección. Los filtros de imagen también son utilizados para realizar realce, así, en [Lalonde et al., 2001] se aplica un filtro de Canny adaptativo para obtener los principales bordes de la imagen como paso previo a la segmentación del DO.

La actuación sobre el espacio de color consiste en seleccionar, combinar o transformar los canales de color a fin de facilitar el proceso de detección (segmentación). Los retinógrafos suelen registrar las imágenes en el espacio de color más común, el RGB. Sin embargo, hay canales que aportan más información que otros para ciertas estructuras, por lo tanto, otro método de preprocesamiento relacionado con el color es el de utilizar solo ciertos canales y desechar el resto. Por ejemplo, a pesar de que el canal verde es el que proporciona mejor contraste [Walter and Klein, 2005], otras estructuras oculares, como el DO, suelen aparecer como una silueta blanca bien definida en el canal rojo [Joshi et al., 2010b, Jelinek et al., 2005]. Sin embargo, existen otros trabajos donde el DO no es discernible en el canal rojo, utilizando entonces el canal verde para realizar la segmentación [Hoover and Goldbaum, 2003, Lalonde et al., 2001, Welfer et al., 2010, Abdel-Ghafar and Morris, 2007, Chrástek et al., 2002]. En otros trabajos se realiza la fusión de estos dos canales de acuerdo a un determinado esquema [Tang et al., 2003, Lu and Lim, 2010a], o incluso la fusión de los tres canales RGB según distintos pesos [Girard et al., 2016]. Algunos estudios, como el que se realiza en [Osareh et al., 2002b], realizan comparativas entre diversos espacios de color para destacar el que

mejor funciona a la hora de detectar el DO. Los canales también pueden normalizarse para abarcar todo el rango posible de valores de intensidad, combinarse en uno solo (conversión RGB a niveles de gris), o incluso transformarse a otros espacios de color como, por ejemplo, el denominado HSI (del inglés, *Hue, Saturation and Intensity*) [Tang et al., 2003, Qureshi et al., 2012].

Por otra parte, durante el proceso de adquisición de la imagen puede aparecer ruido, por lo que es necesario reducirlo. Para ello, se han utilizado distintos tipos de filtros como por ejemplo: filtro Gaussiano [Niemeijer et al., 2009, Welfer et al., 2010], filtro de mediana [Xiong and Li, 2016] o filtro de Wiener [Carmona et al., 2008].

En cuanto al reescalado de imagen, este suele ser bastante utilizado porque aporta varias ventajas: reduce el coste computacional y elimina posibles artefactos ruidosos que solo aparecen a la escala original [Welfer et al., 2010, Giachetti et al., 2011, Morales et al., 2013, Marin et al., 2015, Mittal and Sivaswamy, 2015, Girard et al., 2016].

Por último, la reducción del área de trabajo pretende centrar la atención en la zona donde se encuentran las estructuras de interés, eliminando otras zonas de la imagen no necesarias para el proceso y, por tanto, reduciendo el coste computacional. En general, esta técnica se aplica en aquellos trabajos que parten de la detección de una estructura para posteriormente segmentarla [Carmona et al., 2008, Yu et al., 2012, Dashtbozorg et al., 2015] o como punto de inicio para detectar (segmentar) otras estructuras adicionales [Welfer et al., 2010, Kao et al., 2014, Mittal and Sivaswamy, 2015]. En ambos casos, la estructura inicial detectada se utiliza como referencia para extraer una ventana de tamaño adecuado. Por tanto, la detección previa de una estructura anatómica puede verse también como una etapa de preprocesado de la tarea de detección o segmentación.

Adicionalmente, la detección o segmentación de una estructura, para posteriormente ser eliminada de la escena, puede considerarse también como una etapa de preprocesado. Así, por ejemplo, una técnica muy utilizada en imágenes de retina es el *inpainting* de la red de vasos, consistente en restaurar o realzar dicha estructura. Normalmente, tras el *inpainting*, se suele eliminar dicha red. La eliminación es necesaria ya que en algunos casos podría considerarse un distractor en el proceso de detección (segmentación) de otras estructuras [Osareh

et al., 2002a, Nyul, 2009, Marin et al., 2015, Dashtbozorg et al., 2015]. En [Osareh et al., 2002a], previo a la detección del DO, se aplican operadores morfológicos en distintos espacios de color a fin de detectar la red de vasos y, posteriormente, eliminarla de la escena. Por otra parte, en [Nyul, 2009], se parte de la información de intensidad y características de los vasos, aplicando un umbralizado adaptativo a fin de realizar una detección inicial de la red vascular. Posteriormente, se crea una máscara binaria combinando la detección inicial de la red vascular con el resultado de aplicar un detector de Canny a la imagen inicial. A partir de la máscara binaria se realiza el *inpainting* de los vasos para posteriormente eliminarlos, obteniendo una imagen donde poder estudiar posibles patologías asociadas al DO. Finalmente, en [Dashtbozorg et al., 2015, Marin et al., 2015] se vuelven a aplicar operadores morfológicos y distintos filtrados con el mismo objetivo, detectar y eliminar la red de vasos de la imagen para no entorpecer la segmentación del DO.

## 2.2. Detección de estructuras anatómicas retinianas

Como ya se indicó en la sección 1.1, la detección de estructuras tiene una gran importancia, aportando información valiosa para el proceso de diagnóstico de patologías, utilizándose también como puntos de referencia (*landmarks*) para la detección de otras estructuras y, adicionalmente, como un paso previo para la segmentación de las mismas, favoreciendo así la precisión y reduciendo el coste computacional del proceso de segmentación.

Muchas estrategias han sido utilizadas para detectar las diferentes estructuras anatómicas en imágenes retinianas. En una primera clasificación, los diferentes enfoques se pueden dividir en dos grandes grupos: los que solo utilizan el conocimiento de las propiedades de cada estructura (CRintra-e) para su detección de forma aislada y aquellos que, además, también utilizan el conocimiento relacional entre diferentes estructuras (CRinter-e) para realizar la detección conjunta de dos o más estructuras. La Fig. 2.1 muestra, en forma de árbol, la jerarquía seguida a lo largo de esta sección para introducir los diferentes trabajos de detección de estructuras retinianas.

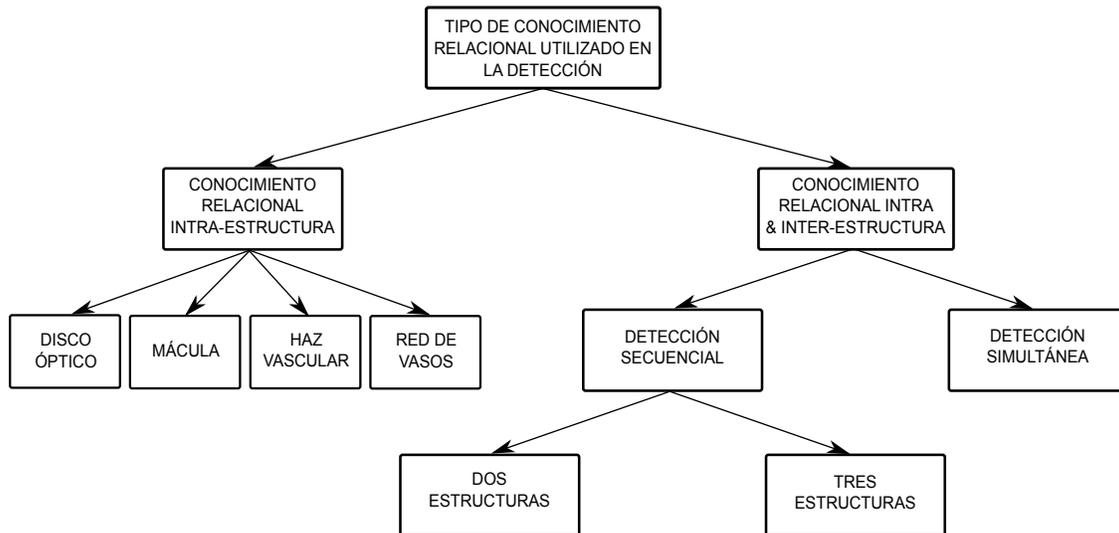


Figura 2.1: Clasificación de los diferentes trabajos de detección de estructuras retinianas.

### 2.2.1. Uso aislado del CR<sub>intra-e</sub>

Son pocos los trabajos que utilizan únicamente el CR<sub>intra-e</sub> para detectar una única estructura. La mayoría de ellos detectan como mínimo dos estructuras haciendo uso de ambos tipos de conocimiento (CR<sub>intra</sub> e inter-e). Esto es debido a la dificultad existente en la detección aislada. Además, resulta más difícil detectar la mácula que el DO (el contorno de la mácula no suele estar tan definido como el del DO). Por último, la posibilidad de aparición de signos patológicos en la imagen en forma de manchas brillantes y/o oscuras (falsos positivos) dificulta todavía más el proceso correcto de detección. Los trabajos presentados en esta sección no hacen uso de las bondades del CR<sub>inter-e</sub> a la hora de aprovechar las relaciones geométricas existentes entre estructuras y poder así facilitar la eliminación de falsos positivos. De otra parte, debido a las características de la red vascular, la cual se distribuye a lo largo y ancho de la retina, no se han utilizado enfoques dirigidos a una aproximación previa de la misma para delimitar una subregión de búsqueda y, posteriormente, segmentarla con más precisión, tal y como sí suele hacerse con el DO o la mácula. Esta circunstancia hace que la mayoría de los trabajos existentes en la literatura relacionados con la red vascular estén dirigidos a la segmentación precisa de esta estructura. Sin embargo, como se verá en la Sección 2.2.2, la detección rápida y aproximada de la red vascular puede ser útil para

facilitar el proceso de detección de otras estructuras. Además, y dado que el HV se podría considerar como una estructura perteneciente a la red vascular, para detectar la primera necesitaríamos la segunda. Por tanto, no tiene mucho sentido realizar la detección del HV utilizando solamente el CR<sub>intra-e</sub> y sí (como veremos más adelante) utilizando ambos tipos de conocimiento.

### Disco óptico

La detección del DO no es una tarea sencilla. Además de la posible variación de forma, tamaño y color, existen algunas complicaciones adicionales a tener en cuenta. Por ejemplo, las marcas brillantes en forma de exudados duros o manchas algodonosas actúan como distractores en el proceso de detección del DO, pudiendo dar lugar a falsos positivos. Los primeros métodos semiautomáticos de detección de estructuras retinianas se basaban en propiedades geométricas de los píxeles de la imagen y sus valores de intensidad [Cox and Wood, 1991]. Otro ejemplo basado en esta idea, lo tenemos en [Sinthanayothin et al., 1999], donde, de una forma totalmente automática, se utiliza la variación de intensidad, el *template matching* y las redes neuronales para la detección de tres estructuras de forma independiente y utilizando el conocimiento propio de cada una de ellas. En particular, el DO se detecta identificando el área con la mayor variación en intensidad de los píxeles adyacentes. El problema de las técnicas basadas en el valor de intensidad de los píxeles de DO viene dado por posibles diferencias en estos valores de luminosidad dependiendo de la imagen considerada, las cuales podrían afectar a la bondad de la detección. A pesar de detectar varias estructuras en el mencionado trabajo, en ningún momento se hace uso del CR<sub>inter-e</sub>.

En [Lalonde et al., 2001] se aplican técnicas de descomposición piramidal y *template matching* en la tarea de detección del DO. En este trabajo, la descomposición piramidal divide la imagen original en un subconjunto de subimágenes por niveles, donde cada nivel corresponde a una resolución cada vez menor. La idea que propone este trabajo es que tanto el DO como el resto de artefactos luminosos parecidos seguirán apareciendo en la imagen a bajas resoluciones. En principio, el DO es la zona más brillante de la retina (junto con posibles exudados y otras lesiones brillantes), por lo que el problema se reduce a la búsqueda

del verdadero positivo en un subconjunto de regiones candidatas formado por varios falsos positivos. Finalmente, se aplica *template matching* basado en la distancia de Hausdorff para seleccionar el DO de entre las regiones candidatas.

En un trabajo más reciente [Aquino et al., 2010], se realiza una detección del DO a partir de tres métodos diferentes, cada uno de los cuales obtiene una posición candidata de esta estructura. Cada método se basa, respectivamente, en: (i) encontrar la zona más brillante de la retina; (ii) la de máxima varianza; y (iii) la de máxima intensidad (tras aplicar un filtro paso-bajo). Por último, se aplica un sistema de votación, que depende de las posiciones relativas las soluciones candidatas, para obtener un punto final como aproximación al centro del DO.

## Mácula

Bajo el paradigma que solo hace uso del CR<sub>intra-e</sub> para la detección de una estructura, la mácula ha recibido menos atención que el DO, probablemente debido al hecho de que la detección de mácula aislada es más difícil de resolver que la detección de DO. Además, la detección de la mácula también se ve afectada por signos visuales asociados a diferentes patologías, ya sea porque oculten la zona de interés parcialmente (debido a manchas brillantes en forma de exudados) o bien porque tengan similar forma y/o tamaño (manchas oscuras o hemorragias). Una de las propiedades maculares más utilizadas en la detección es aquella que establece que, en ausencia de patologías o de ruido, el centro de la mácula corresponde a la región de píxeles de menor intensidad de la retina. Así, en [Sinthanayothin et al., 1999] la mácula se detecta mediante la correlación de coincidencia de la retina con un *template* que representa esta estructura. Dicho *template* es creado utilizando las características típicas de la mácula, esto es, oscuridad de la región y ausencia de vasos. La localización de la máxima correlación entre la intensidad de la imagen y dicho *template*, es escogida como ubicación de la mácula.

En [Petsatodis et al., 2004], se escoge automáticamente una zona rectangular con la acumulación de píxeles de menor intensidad. Posteriormente, se realiza una umbralización a fin de concretar el área de píxeles con menor intensidad de dicha zona. Finalmente, se aplica un

filtro paso-bajo gaussiano para difuminar la zona seleccionada, y se marca el centroide de dicha zona como el centro de mácula.

En [Lu and Lim, 2010a], también se hace uso de la mínima intensidad asociada a la mácula, combinada con el uso de operadores lineales. Así, dado que el nivel de intensidad de los píxeles de mácula crecen gradualmente conforme nos vamos alejando de su centro, la idea es detectar aquel punto donde el cambio de nivel de gris es mayor conforme nos vamos alejando de él. El operador lineal lo que hace básicamente es realizar una transformación lineal entre espacios vectoriales topológicos. En este caso, dado un punto y una serie de líneas trazadas con distintos ángulos y centradas en dicho punto, la información de niveles de gris es transformada en un valor numérico que representa el grado de cambio en la intensidad con la distancia. Este operador es aplicado a toda la imagen, dando finalmente como solución del centro de mácula aquel punto que cumple en mayor grado el operador definido.

### 2.2.2. Uso conjunto de CR<sub>intra</sub> e inter-e

Los trabajos que detectan dos o más estructuras suelen utilizar ambos tipos de conocimiento relacional: intra e inter-e. Concretamente, en esta sección, nos centramos en este tipo de trabajos. Además, la inmensa mayoría de ellos realizan la detección de forma secuencial, es decir, detectando primero una estructura mediante CR<sub>intra-e</sub>, utilizando el resultado de esta como punto de referencia (CR<sub>inter-e</sub>) para detectar la siguiente, y así sucesivamente. De otro lado, la detección simultánea de dos o más estructuras ha sido muy poco explotada en la literatura relacionada.

#### Detección secuencial de dos estructuras

Existen numerosos ejemplos en los que se realiza la detección secuencial de dos de las principales estructuras retinianas. Por ejemplo, la segmentación previa de la red principal de vasos ha sido utilizada como marco de referencia en la detección de DO [Hoover and Goldbaum, 2003, Foracchia et al., 2004, Ying et al., 2007, Youssif et al., 2008, Mahfouz and Fahmy, 2010, Marin et al., 2015, Xiong and Li, 2016], de la mácula [Chutatape, 2006] o del HV [Wel-

fer et al., 2010, Salazar-Gonzalez et al., 2014]. También la detección previa de DO ha sido explotada para restringir la búsqueda de la mácula en una subregión de la imagen original, situada a una distancia y orientación típicas en relación al DO [Sekhar et al., 2008, Welfer et al., 2011, Kao et al., 2014, Mittal and Sivaswamy, 2015]. A continuación, se describe con algo más de detalle las referencias mencionadas.

### *Detección de DO a partir de la red de vasos*

Empecemos por aquellos trabajos en los que se utiliza la segmentación previa de la red vascular como marco de referencia para detectar el DO. Así, por ejemplo, en [Hoover and Goldbaum, 2003], se aplica el nivel de intensidad de la imagen y una estrategia basada en la convergencia difusa de la red de vasos para detectar el DO. La idea subyacente en la convergencia difusa de vasos es detectar la zona de vasos con mayor densidad. Aquí se explotan dos propiedades del DO y de la red de vasos: los vasos principales convergen hacia la zona central del DO y, de otro lado, el DO es la zona más brillante de la imagen.

En [Foracchia et al., 2004] se crea un modelo geométrico de la red de vasos asumiendo la forma parabólica de sus arcadas superior e inferior. Por tanto, el punto de unión de las dos ramas de la parábola determinará el centro aproximado del DO. La limitación de este método depende del grado de contraste de la red de vasos en la imagen.

En [Ying et al., 2007] se usa la red de vasos esqueletonizada para la detección del área papilar. Aquí, en primer lugar, se seleccionan aquellas regiones más brillantes de la imagen (las cuales incluyen el DO y posibles lesiones brillantes) como soluciones candidatas de DO. Dado que la mayor convergencia de vasos se produce en el interior del DO, la idea es quedarse con aquella solución candidata en la que la convergencia de vasos es máxima. Para ello, se aplica un análisis fractal a aquellas zonas candidatas de la imagen esqueletonizada de la red de vasos. La idea del análisis fractal es que un patrón permanece constante cuando se analiza en distintos niveles de escalado. Así, en la zona papilar habrá mayor dimensión fractal que en aquellas zonas donde solo hay lesiones brillantes, las cuales suelen ser zonas avasculares. El análisis es aplicado a una subventana centrada en cada una de las regiones candidatas, escogiendo aquella subventana que devuelva mayor dimensión fractal. Finalmente se aplica un

análisis de histograma escogiendo el centro del DO de entre el 25 % de píxeles más brillantes de la subventana elegida. El método es robusto ya que es capaz de descartar falsos positivos (lesiones brillantes), principales distractores en la búsqueda del DO.

La convergencia de los vasos se explota también en [Youssif et al., 2008] a la hora de detectar el DO. Aquí, inicialmente, se normaliza la luminosidad de la imagen y, además, se mejora el contraste a través del método de ecualización del histograma adaptativo. A continuación, se detectan los vasos de forma aproximada mediante un filtro de borde y, tras varios refinamientos, se obtiene una red de vasos más precisa. Por otra parte, se crea un patrón general de las direcciones que tienen los vasos en la vecindad del DO. El último paso consiste en aplicar un *matched-filter* a la red de vasos con el patrón de direcciones. El punto con mayor respuesta será el centro del DO.

Otra técnica más reciente, utiliza proyecciones unidimensionales de la imagen para realizar la detección del DO basado en la red vascular [Mahfouz and Fahmy, 2010]. En este trabajo, a fin de buscar los vasos, se aplican filtros de borde horizontal y vertical. Este trabajo explota la característica de los vasos por la cual estos emergen del DO hacia arriba o hacia abajo (en el eje vertical) cuando están cerca del centro del DO y, se dirigen hacia la izquierda o la derecha (eje horizontal), cuando están más alejados de dicho centro. La información 2D del resultado de aplicar el filtro de bordes a la imagen se transforma a dos proyecciones de amplitud unidimensionales, una para el *eje-x* y otra para el *eje-y*. En primer lugar, se calcula la posición horizontal del DO (punto de mayor amplitud en el *eje-x*) para luego calcular la posición vertical. El valor más alto en la señal de amplitud de proyección a lo largo de ambas direcciones indicará la posición del DO. El problema de este método es que depende en gran medida de la bondad del proceso de detección horizontal. De forma parecida, en [Xiong and Li, 2016], se utiliza una técnica basada en filtros de borde. Aquí, en primer lugar, a partir de los mapas (horizontal y vertical) obtenidos mediante dichos filtros, se busca puntos candidatos de DO tanto en el plano horizontal como en el vertical. A partir de estos puntos se calcula el área y brillo de las zonas a las cuales pertenecen. Con esta información, se calcula el valor de una función de ajuste. El mayor valor obtenido con esta función da el punto más plausible de pertenecer al DO. Este método es robusto frente a casos patológicos donde aparecen lesiones

brillantes y ante casos donde el DO se encuentra parcialmente visible.

Finalmente, en [Marin et al., 2015], se utiliza la información de la red de vasos (apoyada en el uso de morfología) para obtener el centro del DO. En primer lugar, en este trabajo se reduce el tamaño de las imágenes para que el coste computacional no sea demasiado elevado y, posteriormente, se segmenta la red de vasos. Mediante sucesivas operaciones morfológicas se extrae dicha red del fondo, obteniendo así una imagen difuminada donde, en condiciones normales, la única estructura que permanece es el DO. Aprovechando las características de intensidad propias del DO y del posible nivel de gris residual de los vasos extraídos, se aplica un proceso de umbralización a fin de obtener una ROI del DO. Tras una segunda umbralización se obtienen aquellos píxeles más brillantes de la ROI, cuyo centroide representa el centro del DO.

#### *Detección de mácula a partir de la red de vasos*

La localización de la red de vasos también ha sido utilizada como paso previo para la posterior localización de la mácula. Así, por ejemplo, en [Chutatape, 2006], se identifican los vasos principales utilizando una variante de los contornos activos denominada modelo de forma activo (ASM, del inglés *Active Shape Model*) y, a continuación, el conjunto de puntos obtenido se ajusta mediante una parábola. El ASM parte de un modelo de puntos que representa la estructura a buscar y es creado a partir de un conjunto de entrenamiento. A partir de dicha parábola y conociendo la distancia media entre el punto desde el que emergen los vasos y la mácula, se detecta esta última. Dado que el centro del DO y el centro del HV están muy próximos, dicha distancia es prácticamente la misma que la bien conocida distancia media existente entre los centros de DO y mácula.

#### *Detección del Haz Vasculare a partir de la red de vasos*

La detección de la red de vasos es un paso necesario para obtener el punto de unión de ambas arcadas (centro del HV). Posteriormente, este punto puede ser utilizado para detectar otras estructuras adicionales como puede ser el DO.

En [Welfer et al., 2010], se emplea un algoritmo iterativo basado en morfología para ob-

tener una versión simplificada y eskeletonizada de las arcadas principales. Posteriormente, se calcula el centroide de dichas arcadas. La intersección de la recta horizontal que pasa por dicho punto y las arcadas, indicará con alta probabilidad el punto de unión de ambas arcadas o, en su defecto, un punto situado dentro del área papilar. En [Salazar-Gonzalez et al., 2014], la red de vasos se obtiene mediante la denominada técnica *graph-cut*, consistente en representar cada píxel de la imagen como el nodo de un grafo y, a partir de dicha información, realizar una división de fronteras por similitud de intensidad. Partiendo de la red de vasos, se inspiran en el trabajo de [Welfer et al., 2010] para detectar el punto de unión de las arcadas superior e inferior.

### *Detección de mácula a partir del DO*

Por último, describimos algunos de los trabajos relevantes en los que se utiliza la detección previa del DO como punto de referencia para facilitar la detección de la mácula. Así, por ejemplo, en [Sekhar et al., 2008], se parte de una etapa de preprocesado consistente en la eliminación de la red vascular mediante una operación morfológica de clausura con un elemento estructurante más grande que el grosor del mayor ancho de la red de vasos. Como resultado, se obtiene una imagen difuminada, pero sin la existencia de vasos. Tras esto, se aplica la transformada circular de Hough a fin de detectar la forma circular del DO. Posteriormente y utilizando la relación espacial conocida entre DO y mácula (los centros de mácula y DO se encuentra a una distancia media de 2.5 veces el diámetro del DO [Sinthanayothin et al., 1999]), se detecta la zona macular. Finalmente se realiza una umbralización para obtener el centro de la mácula. De forma similar, en [Welfer et al., 2011], también se utiliza morfología matemática y las relaciones geométricas conocidas entre mácula y DO. Así, en primer lugar, haciendo uso del CRinter-e relativo a la distancia típica DO-mácula, y conociendo la posición y diámetro del DO (obtenido mediante un método morfológico comentado en [Welfer et al., 2010]), se extrae una región de interés (ROI, del inglés *Region Of Interest*) en la cual se lleva a cabo la búsqueda de la mácula. A continuación, mediante la aplicación de diferentes operadores morfológicos de apertura y clausura, se eliminan aquellos patrones visuales de patología (exudados, microaneurismas y hemorragias) que pudieran suponer un distractor en

la detección. Seguidamente, se binariza la ROI para seleccionar aquellas regiones candidatas de pertenecer a la mácula. Después, estos candidatos son filtrados haciendo uso de la relación conocida de que el segmento imaginario que une DO y mácula forma un ángulo típico con el eje horizontal [Williams and Wilkinson, 1992]. Finalmente, de los candidatos no filtrados del paso anterior, se escoge como mácula aquel con el nivel de gris más bajo.

De otro lado, en [Kao et al., 2014], se hace uso de *template matching* y del CRinter-e para el proceso de detección de DO y mácula. Aquí, el DO se aproxima mediante *template matching* y, haciendo uso del CRinter-e relativo a la distancia típica DO-mácula, se selecciona una ROI libre de vasos en la que, también, mediante *template matching*, se realiza la búsqueda de la mácula.

Finalmente, en [Mittal and Sivaswamy, 2015], se hace uso de un método basado en lo que en castellano podría denominarse patrón de movimiento generalizado (GMP, del inglés *Generalized Motion Pattern*). Este método se creó inicialmente para detectar anomalías en mamografías e imágenes de retina. Se basa en la inducción de movimiento a una imagen para suavizar la información no deseada y, al mismo tiempo, realzar las estructuras de interés. La idea es que a partir de un punto fijo (pivote) se realizan distintas rotaciones de la imagen, obteniendo como resultado lo que se denomina *mapa de interferencias*, que combina la información de movimiento con la imagen original. El resultado de este proceso supone un difuminado de las regiones luminosas más pequeñas y el realzado de las regiones grandes y luminosas como, por ejemplo, el DO. Así, en primer lugar, aplicando esta técnica se detecta el DO, y posteriormente, utilizando CRinter-e relativo a la distancia DO-mácula, se selecciona una ROI en la que acotar la búsqueda de la mácula. Finalmente, utilizando técnicas de umbralización se calcula el centro de la mácula.

### **Detección secuencial de tres estructuras**

Siguiendo la tendencia de los enfoques mencionados anteriormente, también existen trabajos de investigación en los que se ha propuesto el uso del CRintra e inter-e para detectar de forma secuencial tres estructuras retinianas diferentes [Gagnon et al., 2001, Li and Chutatape, 2004, Tobin et al., 2007, Fleming et al., 2007, Niemeijer et al., 2009, Yu et al., 2011, Giachetti

et al., 2013].

Así, en [Gagnon et al., 2001], se utiliza el DO como punto inicial para la detección de mácula y para la extracción del árbol vascular. El DO se aproxima, en primera instancia, a partir de una descomposición piramidal, usando la propiedad de que la zona de píxeles más brillantes (DO) ha de mantenerse a distintas escalas de imagen. Acto seguido, para mejorar la detección del DO, se aplica un filtro de bordes de Canny y se aplica un *template matching* basado en la distancia de Hausdorff. A continuación, se busca la mácula como el píxel más oscuro situado a la distancia típica DO-mácula. La red de vasos se detecta partiendo del centro del DO y mediante un algoritmo de *tracking* va siguiendo el trazado del vaso hasta que se encuentra una bifurcación o el vaso finaliza. Si ocurre lo primero, guarda la posición de la bifurcación y continúa por uno de los caminos posibles. Este proceso se repite hasta encontrar el final del vaso recorrido. Seguidamente, el recorrido continúa en el camino sin recorrer de la última bifurcación guardada. Este proceso se repite hasta recorrer toda la red.

En [Li and Chutatape, 2004], en primer lugar, se detecta y segmenta el DO. Para ello, se utiliza la técnica de análisis de componentes principales (PCA, del inglés *Principal Component Analysis*). El objetivo del PCA es obtener una nueva imagen en nivel de gris que mantenga las principales características de la imagen original, pero reduciendo la dimensionalidad. A continuación, partiendo de la posición del DO y creando un modelo de forma activa para la red de vasos, se detecta la red de vasos principales. Por último, y haciendo uso de la distancia típica DO-mácula y de la posición de la mácula con respecto al DO y los vasos, se detecta esta última estructura.

En [Tobin et al., 2007], se detecta el DO a partir de la segmentación de la vasculatura y teniendo en cuenta que los vasos más cercanos al DO son los más densos, los más gruesos y los que están orientados verticalmente. Posteriormente, a partir del DO y de un modelo geométrico de la vasculatura, se detecta el rafe horizontal. Finalmente, a partir de este último y de la distancia típica DO-mácula se detecta la mácula.

La forma elíptica de las arcadas principales de la red de vasos también ha sido utilizada para detectar el DO y la mácula [Fleming et al., 2007]. La forma elíptica de los principales vasos sanguíneos de la retina se utiliza para obtener localizaciones aproximadas para el DO

y la mácula, que se refinan sobre la base del borde circular de la primera y el oscurecimiento local en la última. Así, en primer lugar se realizan mejoras de contraste encaminadas a realzar la red de vasos. Acto seguido se aplica una transformada de Hough para encontrar la forma semi-elíptica de la arcada temporal, buscando el conjunto de semi-elipses que mejor se ajusta a dicha arcada. Se utilizan múltiples semi-elipses para que el proceso de búsqueda sea más robusto. A partir de los centros del conjunto de semi-elipses, se declara una ROI como zona aproximada del DO. Seguidamente, se aplica la transformada circular de Hough para encontrar esta estructura en la ROI. Finalmente, y volviendo a utilizar las propiedades conocidas entre las estructuras (vasos-DO-mácula) se localiza la región macular. Para ello, se traza la línea horizontal que parte del centro del DO y que divide las arcadas por la mitad, detectando la zona macular en la parte más oscura situada a la distancia típica DO-mácula.

En [Niemeijer et al., 2009], se utiliza un regresor no paramétrico (no hace suposiciones sobre la distribución que siguen los datos), basado en el conocido método de los  $k$ -vecinos más cercanos ( $k$ -NN, del inglés *k-Nearest Neighbors algorithm*) para detectar el DO y mácula, asumiendo como conocida la vasculatura retiniana, es decir, en este trabajo no segmenta los vasos, siendo la red vascular una entrada al método propuesto. El regresor  $k$ -NN es utilizado para predecir la distancia de cualquier punto de la imagen a las estructuras de interés, a partir de la información proporcionada por la segmentación de la red vascular y de la propia imagen. El regresor se entrena con un conjunto de imágenes donde la posición del DO y mácula son conocidas. Las características utilizadas se obtienen a partir de dos *templates* circulares, uno para el DO y otro para la mácula, aplicados tanto a la imagen como a la red de vasos. Concretamente, las características que se extraen de la red de vasos son comunes a ambos *templates*: número de vasos, grosor medio, desviación estándar del grosor medio, máxima anchura, densidad media, etc. De otro lado, las características que se extraen de la imagen dependen del tipo de *template*: en el caso del *template* de DO, se extrae la media y la desviación estándar de la intensidad de los píxeles abarcados por dicho *template*; en el caso de mácula, se hace lo mismo, pero divide dichas medidas en dos subconjuntos: medidas de la parte más exterior y de la parte más interior del *template*. Tras aplicar el modelo de regresor aprendido con el *template* de DO a todos los píxeles de una imagen, el punto con

menor distancia es seleccionado como solución del centro del DO. La subregión de búsqueda para la mácula se define en base a la relación conocida de distancia típica entre DO-mácula. Es en dicha subregión de imagen donde se aplicará el regresor con el *template* circular de mácula, para finalmente devolver la posición del centro de mácula.

Por otra parte, en [Yu et al., 2011], se hace uso del *template matching* para la detección del DO y la mácula, pero, con el fin de reducir los falsos positivos debidos a la existencia de manchas brillantes (ruido o signos de patología), también se explota las características de la red de vasos que atraviesan el DO. Así, en primer lugar, mediante *template matching* se selecciona un conjunto de candidatos a DO. A fin de descartar falsos positivos (exudados y otros artefactos brillantes), al conjunto de candidatos se aplica un filtro de emparejamiento direccional (DMF, del inglés *Directional Matched Filtering*). El filtro se configura para detectar aquellos candidatos que son atravesados por vasos. Por último, la mácula se detecta a partir de la posición del DO y la relación de distancia típica DO-mácula.

Finalmente, en [Giachetti et al., 2013], se utiliza la denominada transformada de simetría radial rápida (FRS, del inglés *Fast Radial Symmetry*) para detectar el DO y la mácula. En particular, se utiliza la propiedad de que los píxeles de DO suelen ser brillantes y, los de mácula, oscuros, teniendo ambas estructuras cierta simetría por separado. Por tanto, al aplicar la transformada FRS a la imagen de retina, se obtienen dos mapas de respuesta que indican la simetría que encuentra en dicha imagen con respecto a un *template* circular. De una parte, se obtiene un mapa de gradiente positivo, es decir, un mapa que indica respuestas altas al *template* dado. De otra parte, se obtiene un mapa de gradiente negativo indicando respuestas bajas al mismo *template*. Por tanto, el primero se puede utilizar para estimar la posición del DO y, el segundo, para estimar la posición de la mácula. Por otra parte, también se hace uso de la información de densidad de los vasos: el DO suele ser atravesado por la red de vasos, mientras que la zona macular suele estar libre de vasos. Por tanto, se utiliza la red de vasos para descartar falsos positivos (hemorragias, microaneurismas o exudados).

Todos los trabajos de detección descritos que realizan una detección basada en la detección secuencial de una o más estructuras previas tienen la ventaja de reducir la posibilidad de obtener un falso positivo como solución, ya que usan el conocimiento relacional existente

entre las diferentes estructuras. Sin embargo, también tienen el inconveniente de que un fallo en la detección de al menos una de las estructuras usadas como referencia provocará con alta probabilidad un fallo en la búsqueda de la estructura que se quiere detectar. Para minimizar este último inconveniente, el conocimiento relacional debería usarse en el marco de una detección simultánea. De esta forma, al realizar la búsqueda de forma simultánea, el conocimiento relacional no presupone la existencia de una estructura detectada de antemano. Nótese que una detección basada en la búsqueda sucesiva de diferentes estructuras, no podrá usar el conocimiento relacional en la detección de la primera estructura, es decir, solo podrá usar el conocimiento relacional entre las estructuras  $i$  e  $(i + 1)$ , siendo  $i \geq 2$ .

### DetECCIÓN SIMULTÁNEA

Hasta donde los autores conocen, solo se han podido encontrar tres trabajos en el que se aborda la detección simultánea de dos o más estructuras, en concreto, del DO y la mácula [Qureshi et al., 2012, Dashtbozorg et al., 2016, Al-Bander et al., 2018].

En [Qureshi et al., 2012] se aplican técnicas de diferentes autores para la detección de DO. Lo mismo se hace para la detección de mácula. Cada una de estas técnicas produce una solución candidata para el DO o para la mácula. Posteriormente, se aplica una estrategia para elegir la mejor pareja DO-mácula de entre las soluciones candidatas. Aquí, las técnicas utilizadas para detectar el DO son cinco: descomposición piramidal, detección de bordes, filtro de entropía, transformada circular de Hough y vectores de características de la red de vasos. Las técnicas aplicadas para la detección de mácula son otras cinco: basada en intensidad, relación espacial mácula y DO, relación espacial mácula y arcada temporal, operadores morfológicos y transformación *Watershed*. El método tiene una fase de entrenamiento y otra de test. Una vez obtenidos los conjuntos candidatos de DO y mácula, se aplica el mismo método en dos pasos a cada uno de dichos conjuntos: (i) las coordenadas de cada solución son corregidas basándose en el desplazamiento medio del conjunto de soluciones candidatas con respecto a la posición real de la estructura (información extraída del conjunto de entrenamiento); y (ii) a fin de localizar agrupaciones de soluciones candidatas, se aplica un proceso de detección basado en lo que se denominan zonas activas (*hotspots*), es decir, se aplica un *template* cir-

cular centrado en cada uno de las soluciones candidatas y se realiza un recuento del resto de soluciones candidatas que caen dentro de dicho círculo. Aquellos círculos con mayor número de soluciones candidatas dentro de su perímetro son designados como regiones activas. El número de regiones activas por cada imagen no es conocido a priori, dependiendo en cada caso de la distribución espacial de las soluciones candidatas en dicha imagen. Finalmente, las regiones activas pertenecientes a cada conjunto (mácula y DO) son combinados en parejas mácula-DO. Aquella pareja que satisface en mayor medida las restricciones geométricas (basadas en distancias y ángulos) es escogida como la solución final mácula-DO.

En [Dashtbozorg et al., 2016], se utilizan los denominados filtros superelípticos (SEF, del inglés *Super-Elliptical Filter*) para detectar el DO y la mácula. A su vez, los SEFs se basan en los denominados filtros de índice de convergencia (CIF, del inglés *Convergency Index Filter*), los cuales son adecuados para la detección de formas convexas, independientemente de su contraste con respecto al fondo. En concreto, los CIFs evalúan el grado de convergencia de los vectores de gradiente dentro de un área local (región de interés de búsqueda) con respecto a un píxel de interés. Dicho píxel es el punto de aplicación del filtro, que normalmente será el centro de la estructura a detectar. Es decir, se está evaluando el grado de variación de los píxeles en un área con respecto a un punto central. Por tanto, un SEF se define como un tipo de CIF, pero configurado para detectar formas convexas semi-elípticas, y por tanto, adecuado para la búsqueda del DO y la mácula. La máxima respuesta del filtro devolverá la localización de la estructura buscada. El método final se denomina PSEF (del inglés *Paired SEF*), el cual combina dos filtros SEF configurados para cada una de las estructuras, haciendo uso además de la distancia y posición relativa entre el DO y la mácula. Esta combinación de filtros y CRinter-e realiza la detección conjunta y simultánea de ambas estructuras.

En [Al-Bander et al., 2018], se utilizan redes neuronales convolucionales para realizar la detección simultánea. En este trabajo se hace uso de una única red neuronal convolucional de quince capas (en las doce primeras, las neuronas se distribuyen espacialmente en 2D y, en las tres últimas, en 1D) para, de forma simultánea, detectar el DO y la mácula. Las redes neuronales convolucionales están compuestas por capas con distinto cometido, siendo una versión mejorada del perceptrón multicapa, pero especialmente adecuadas para clasificación

de objetos en imágenes. Se parte de una capa de entrada ( $256 \times 256$ ), seguida por un conjunto de capas distribuidas en cuatro bloques bien diferenciados: 1) un primer bloque donde aprende 8 filtros, formado por cuatro capas convolucionales seguida de una operación de *dropout*; 2) un segundo bloque donde aprende 16 filtros, formado por tres capas convolucionales, una de muestreo (*pooling*) y otra operación *dropout*; 3) un tercer bloque donde se aprenden 32 filtros, formado por tres capas convolucionales más una de *pooling*; y 4) un último bloque compuesto por dos capas de neuronas ordenadas unidimensionalmente, intercalando entre ambas otra operación de *dropout*. Finalmente la última capa unidimensional se encuentra conectada a una capa de salida de cuatro neuronas. La capa convolucional está compuesta por neuronas convolucionales que, como su propio nombre indica, aprenden/aplican diferentes filtros convolucionales. La capa de muestreo (*pooling*) realiza un submuestreo del resultado obtenido en la capa convolucional inmediatamente anterior. La operación de *dropout* consiste en desconectar un porcentaje de las neuronas en cada iteración del entrenamiento, mejorando así la generalización de la red. El tiempo de ejecución empleado por este método es bastante bajo (0.007s), no obstante, el tiempo de entrenamiento no está especificado en el trabajo. Por tanto, el número de capas y tamaño inicial de muestreo hace pensar que el tiempo de entrenamiento pudiera ser bastante elevado. A pesar de esto, como ventaja cabe decir que es la propia red neuronal la que usa el conocimiento CRintra e inter-e de forma automática durante la fase de entrenamiento para construir la red.

Por último, con el fin de resumir cada uno de los trabajos descritos anteriormente, en la Tabla 2.1 se muestra el valor de diferentes descriptores que permiten la clasificación de dichos trabajos desde distintos puntos de vista: método o métodos principales utilizados para la detección; tipo de estructuras detectadas; carácter independiente, secuencial o simultáneo de la detección; estructura requerida como entrada (y obtenida mediante un método descrito en un paper diferente); tipo de conocimiento utilizado; y tiempo empleado. Los trabajos se presenta en la tabla ordenados cronológicamente, de más antiguo a más reciente.

A partir del contenido de la mencionada tabla, se pueden realizar las siguientes observaciones: 1) la detección de estructuras aisladas basada en el uso único de CRintra-e solo se aplicó en los inicios, para posteriormente dar paso a estrategias basadas en la detección de

## 2.2. DETECCIÓN DE ESTRUCTURAS ANATÓMICAS RETINIANAS

---

dos o más estructuras y en las que se hace uso tanto del CRintra como del CRinter-e; 2) cuando se realiza detección secuencial de dos estructuras, las estrategias de detección del DO se apoyan mayoritariamente en una localización previa de la red vascular; 3) la mayoría de las estrategias de detección de mácula se basan en una detección previa del DO; 4) en relación a la detección secuencial de tres estructuras, la mácula sigue siendo la última estructura detectada y su localización se apoya en una detección previa de DO y vasos; y 5) la detección simultánea de más de una estructura ha sido escasamente abordada en la literatura.

Existen también en la literatura varias referencias dedicadas a recopilar el estado del arte de las distintas técnicas usadas en la detección de las diferentes estructuras retinianas: DO [Mahesh, 2013, Raja and Ravichandran, 2014, Besenczi et al., 2016, Mary et al., 2016], mácula [Besenczi et al., 2016] y red de vasos [Mahesh, 2013, Besenczi et al., 2016]. Sin embargo, en ninguna de las referencias mencionadas se analiza la tarea de detección desde el enfoque aquí descrito, es decir, desde el punto de vista del tipo de conocimiento usado (CRintra o CRinter-e) o de la naturaleza del proceso de detección realizado (una o más estructuras y, en este último caso, si el proceso de detección es secuencial o simultáneo).

Tabla 2.1: Categorización de diferentes trabajos de detección atendiendo a varios descriptores

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras detectadas	Tipo de detección	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Cox and Wood, 1991]	Propiedades geométricas e intensidad de píxeles	DO	Independ.	-	CRintra-e	120
[Sinthanayothin et al., 1999]	Variación de intensidad de píxeles, <i>Template Matching</i> y redes neuronales	DO, mácula, vasos	Independ.	-	CRintra-e	-
[Lalonde et al., 2001]	Descomposición piramidal y <i>Template Matching</i>	DO	Independ.	-	CRintra-e	1.6 ± 0.3
[Gagnon et al., 2001]	Descomposición piramidal, filtros de borde y <i>Template Matching</i>	DO → (mácula, vasos)	Secuenc.	-	CRintra/inter-e	-
[Hoover and Goldbaum, 2003]	Convergencia difusa de la red de vasos	Vasos → DO	Secuenc.	Vasos	CRintra/inter-e	-
[Patsotidis et al., 2004]	Intensidad de píxeles	DO, mácula, vasos	Independ.	-	CRintra-e	25
[Foracchia et al., 2004]	Propiedades geométricas	Vasos → DO	Secuenc.	Vasos	CRintra/inter-e	120
[Li and Chutatape, 2004]	Análisis PCA ( <i>Principal Component Analysis</i> ) y ASM ( <i>Active Shape Model</i> )	DO → vasos → mácula	Secuenc.	-	CRintra/inter-e	-
[Chutatape, 2006]	ASM ( <i>Active Shape Model</i> )	Vasos → mácula	Secuenc.	-	CRintra/inter-e	-
[Ying et al., 2007]	Análisis fractal e intensidad de píxeles	Vasos → DO	Secuenc.	Vasos	CRintra/inter-e	10
[Fleming et al., 2007]	Propiedades geométricas y transformada circular de Hough	Vasos → DO → mácula	Secuenc.	Vasos	CRintra/inter-e	120
[Tobin et al., 2007]	Propiedades geométricas	Vasos → DO → mácula	Secuenc.	-	CRintra/inter-e	-

Tabla 2.1: Categorización de diferentes trabajos de detección atendiendo a varios descriptores (Cont.)

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras detectadas	Tipo de detección	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Yousif et al., 2008]	Propiedades geométricas e intensidad de píxeles	Vasos→DO	Secuenc.	-	CRintra/inter-e	210
[Sekhar et al., 2008]	Propiedades geométricas y transformada circular de Hough	DO→mácula	Secuenc.	-	CRintra/inter-e	-
[Niemeijer et al., 2009]	Clasificador k-NN	Vasos→DO→mácula	Secuenc.	Vasos	CRintra/inter-e	27.6
[Aquino et al., 2010]	Filtros de imagen e intensidad de píxeles	DO	Independ.	-	CRintra-e	1.67 ± 0.14
[Mahfouz and Fahmy, 2010]	Filtros de borde y proyección unidimensional	Vasos→DO	Secuenc.	-	CRintra/inter-e	0.65
[Welfer et al., 2010]	Conocimiento morfológico de la red de vasos	Vasos→HV	Secuenc.	-	CRintra/inter-e	-
[Lu and Lim, 2010a]	Operadores lineales	Mácula	Independ..	-	CRintra-e	-
[Yu et al., 2011]	<i>Template Matching</i>	Vasos→ DO→mácula	Secuenc.	-	CRintra/inter-e	3.9
[Welfer et al., 2011]	Morfología matemática y relaciones geométricas	DO→mácula	Secuenc.	-	CRintra/inter-e	-
[Qureshi et al., 2012]	Propiedades geométricas, intensidad de píxeles, operadores morfológicos, filtros de borde, etc	DO y mácula	Simult.	-	CRintra/inter-e	-
[Giachetti et al., 2013]	FRS ( <i>Fast Radial Symmetry</i> )	Vasos→(DO,mácula)	Secuenc.	-	CRintra/inter-e	5
[Kao et al., 2014]	<i>Template Matching</i>	DO→mácula	Secuenc.	-	CRintra/inter-e	[2.6, 6.9]

Tabla 2.1: Categorización de diferentes trabajos de detección atendiendo a varios descriptores (Cont.)

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras detectadas	Tipo de detección	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Mittal and Sivaswamy, 2015]	GMP ( <i>Generalized Motion Pattern</i> )	DO→mácula	Secuenc.	-	CRintra/inter-e	[120,180]
[Marin et al., 2015]	Operaciones morfológicas e intensidad de píxeles	Vasos→DO	Secuenc.	Vasos	CRintra/inter-e	5.42 ± 0.98
[Salazar-Gonzalez et al., 2014]	<i>Graph-cut</i>	Vasos→HV	Secuenc.	-	CRintra/inter-e	-
[Dashbozorg et al., 2016]	SBF ( <i>Super Elliptical Filter</i> )	DO y mácula	Simult.	-	CRintra/inter-e	2.4
[Xiong and Li, 2016]	Filtros de borde e intensidad de píxeles	Vasos→DO	Secuenc.	-	CRintra/inter-e	-
[Al-Bander et al., 2018]	Red Neuronal Convolucional	DO y mácula	Simult.	-	CRintra/inter-e	0.007

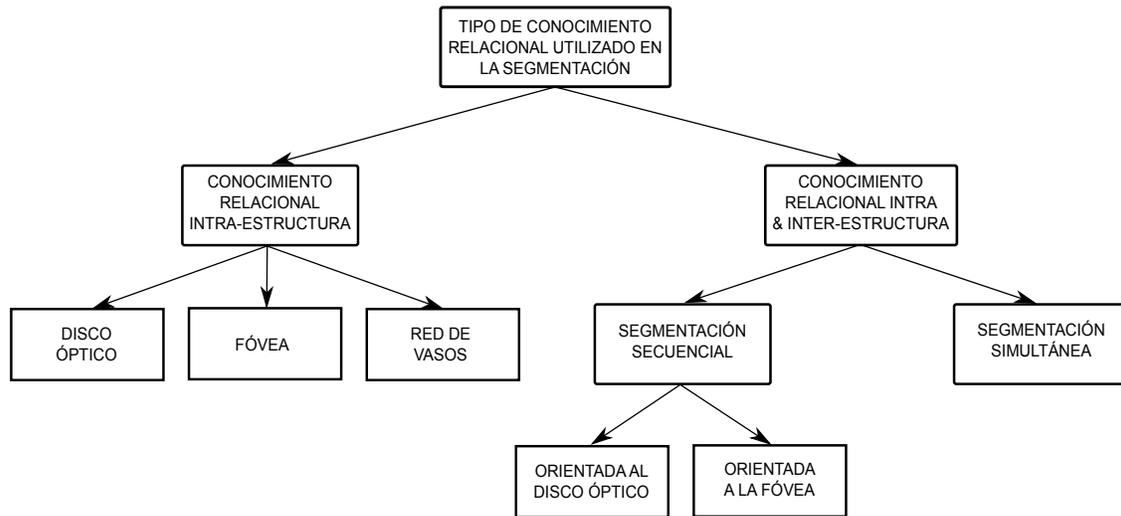


Figura 2.2: Clasificación de los diferentes trabajos de segmentación de estructuras retinianas.

## 2.3. Segmentación de estructuras anatómicas retinianas

Como se comentó en la sección 1.1, la segmentación de estructuras es, si cabe, más importante que la detección. Su importancia radica en el hecho de que para diagnosticar ciertas enfermedades, no basta solo con conocer la localización aproximada de la estructura o estructuras involucradas, es decir, el resultado de detección, sino que es necesario conocer la forma y límites de la misma y, en algunos casos, la evolución temporal de dichas características. La Fig. 2.2 muestra, en forma de árbol, la jerarquía seguida a lo largo de esta sección para introducir los diferentes trabajos de segmentación de estructuras retinianas.

### 2.3.1. Uso aislado de $CR_{intra-e}$

La mayoría de trabajos que solo usan  $CR_{intra-e}$  para la segmentación de una estructura suelen apoyarse en una detección previa de la misma para, posteriormente, refinar su búsqueda en una vecindad (subventana) de dicha aproximación. Como se verá a continuación, esta estrategia es bastante usada en la segmentación del DO. En cambio, las aproximaciones utilizadas para la segmentación de la red vascular solo suelen apoyarse en el uso del  $CR_{intra-e}$ . Dado que esta estructura se distribuye a lo largo y ancho de la retina, no tiene sentido realizar

una aproximación previa de la misma para delimitar una subregión en la que, posteriormente, segmentar dicha estructura.

### Disco óptico

La segmentación precisa del DO no es una tarea trivial debido a la existencia de numerosos distractores. Así, por ejemplo, el contraste asociado al perímetro del DO no suele ser constante debido a la existencia de vasos que oscurecen parcialmente áreas del anillo, provocando la aparición de sombras que dificultan el seguimiento visual de dicho perímetro. Otros distractores están relacionados con la existencia de diferentes patologías como, por ejemplo, la existencia de atrofas peripapilares que pueden producir áreas brillantes alrededor del DO, difuminando o distorsionando su forma.

Existen pocas aproximaciones en las que la segmentación del DO se hace solo y exclusivamente utilizando CR<sub>intra-e</sub> (sin apoyarse en la detección previa). Por ejemplo, la transformada circular de Hough se ha empleado exitosamente en la segmentación del DO [Abdel-Ghafar and Morris, 2007, Zhu et al., 2010, Aquino et al., 2010]. Esta transformada permite encontrar cualquier tipo de figura que pueda ser expresada matemáticamente, como pueden ser rectas, circunferencias y elipses. El problema de esta técnica reside en que su eficacia es altamente dependiente de la calidad de la imagen de entrada, dado que se han de detectar los bordes de la imagen como paso previo a la segmentación. Por tanto, el método puede no trabajar adecuadamente en imágenes que contengan ruido. Por ejemplo, en [Aquino et al., 2010], se realiza un procesado morfológico previo destinado a eliminar los vasos, cuyos contornos podrían actuar como posibles distractores del contorno del DO. Seguidamente, se aplican detectores de bordes y otras técnicas morfológicas para obtener una imagen binaria de contornos parciales candidatos del DO y, finalmente, se usa la transformada circular de Hough para obtener una aproximación circular del contorno del DO. Toda esta operativa se aplica a una subventana centrada en un punto interno del DO obtenido a partir de una etapa previa de detección de esta estructura.

La mayoría de los trabajos de investigación que solo usan CR<sub>intra-e</sub> para la segmentación del DO, se apoyan en la detección previa de esta estructura para delimitar la zona de búsqueda

[Tang et al., 2003, Osareh et al., 2002a, Lowell et al., 2004, Tang et al., 2006, Joshi et al., 2010a, Kim et al., 2007, Aquino et al., 2010, Chrástek et al., 2002, Morales et al., 2013, Dashtbozorg et al., 2015, Carmona et al., 2008, Novo et al., 2009, Lu and Lim, 2010b, Siddalingaswamy and Prabhu, 2010, Yu et al., 2012]. Así, en [Tang et al., 2003], se combinan múltiples imágenes de entrada para componer una imagen resultante (*Image Fusion*) a partir de la cual segmentar el DO. En primer lugar, se detecta un centro aproximado de la estructura mediante un sencillo proceso de umbralización. La idea es tomar los píxeles más brillantes de la imagen. Una vez que la zona de interés ha sido seleccionada, se extraen distintas máscaras centradas en dicha zona. Estas máscaras corresponden a distintos procesos de umbralización aplicados a los canales rojo y verde del modelo RGB, a la imagen varianza y al canal I de la imagen (resultado de transformar la imagen RGB original al espacio de color HSI). Todas estas máscaras se fusionan en una sola, dando una primera aproximación del contorno papilar. Finalmente, se aplica un posprocesado para obtener un contorno más preciso y suave. Este método tiene el problema de falta de robustez frente a lesiones brillantes.

Otras aproximaciones hacen uso de los denominados contornos activos (*snakes*). En los primeros trabajos, [Morris and Donnison, 1999, Mendels et al., 1999, Lee and Brady, 1991], se requería de una etapa de inicialización que implicaba elegir manualmente una curva cercana a la posición del DO. En trabajos más recientes [Osareh et al., 2002a, Lowell et al., 2004, Siddalingaswamy and Prabhu, 2010], la etapa de inicialización se realiza de forma automática, siendo su objetivo el de detectar un punto interno al DO, el cual se toma como referencia para construir una curva cerrada exterior al área papilar. En [Tang et al., 2006, Joshi et al., 2010a], se utiliza una variante de contornos activos basada en regiones y que es conocida como método de Chan&Vese [Chan and Vese, 2001]. En este caso, se obtienen buenos resultados de segmentación de DO sin requerir preprocesamiento ni una etapa previa de eliminación de vasos, que sí requieren otras aproximaciones basadas en contornos activos. No obstante, dado que solamente se utiliza el canal rojo, la aplicación del método en aquellas imágenes con la región papilar saturada o mal iluminada en dicho canal puede dar lugar a una mala segmentación.

El estimador matemático RANSAC (*RANdom SAmple Consensus*) [Fischler and Bolles,

1981] también se ha utilizado en la tarea de segmentación papilar. Básicamente, consiste en un método iterativo para estimar los parámetros de un modelo matemático aplicado a un conjunto de datos que puede contener valores “atípicos”. Los valores normales se ajustarán al modelo matemático y los valores “atípicos” corresponderán a errores de medición o a ruido. Así, en [Kim et al., 2007], se parte de una detección aproximada del centro del DO para, posteriormente, transformar la zona del área papilar a coordenadas polares, tomando como referencia dicho centro. La imagen se binariza posteriormente, seleccionando el primer píxel por columna situado por encima de  $1/3$  del ancho de la imagen (se pretende eliminar ruido cercano al centro del DO). El contorno papilar se aproxima mediante una curva definida por un polinomio de cuarto grado, el cual se inicializa aleatoriamente. A continuación, se define un pequeño umbral de distancia, de tal forma que los puntos cercanos al umbral se etiquetan como “típicos” y, el resto, como “atípicos”. El estimador RANSAC va tomando puntos aleatorios (“típicos” y/o “atípicos”), resolviendo el polinomio y actualizando las etiquetas de los puntos. El contorno papilar será aquel polinomio que maximice el número de puntos “típicos” situados por debajo del umbral establecido. El inconveniente de esta aproximación es que resulta altamente dependiente del grado de exactitud con el que se detecta el centro del DO durante la etapa de inicialización. Una mala aproximación del centro provocará que la transformación a coordenadas polares, que utiliza dicho centro como referencia, altere la forma esperada del contorno de DO.

El análisis PCA también se ha utilizado para segmentar el DO. Así, en [Morales et al., 2013], se aplica PCA a la imagen de entrada para obtener distintas versiones de la misma en nivel de gris. Acto seguido, se aplica una etapa de preprocesado basada en una técnica de restauración de imagen (*inpainting*) para realzar y posteriormente eliminar la red de vasos, ya que supone un distractor para el DO. Posteriormente, se aplica una serie de operadores morfológicos de apertura y cierre con el fin de realzar las partes oscuras y brillantes. Como punto de partida para la segmentación, se calcula el centroide de la imagen de gris obtenida. A partir de dicho punto se aplica una transformación *Watershed*, dando como resultado un conjunto de regiones con diferentes niveles de gris. Las regiones pertenecientes al contorno del DO serán regiones brillantes junto a regiones oscuras (vasos y resto de retina). Por tanto, con

el objeto de aislar las regiones brillantes, se aplica el operador morfológico *close-hole* (este operador resalta un primer conjunto de puntos conectados, denominado *hole*, rodeado por otro conjunto de puntos conectados cuyos niveles de intensidad son mayores que el nivel medio de intensidad del *hole*). La zona restante (zona brillante aislada por el operador anterior) se umbraliza para obtener una versión tosca del área papilar. Por último, al resultado del paso anterior se le aplica una transformación circular de Hough que marca el límite de la zona papilar. En relación a esta última transformación, cabe decir que también se ha utilizado con éxito en otros trabajos [Aquino et al., 2010, Chrástek et al., 2002, Marin et al., 2015] en los que, tras una detección previa del DO y distintas etapas de preprocesado, se aplica dicha transformación para extraer el contorno papilar.

El uso de diferentes filtros de imagen también ha resultado efectivo en la tarea de segmentación del DO. Así, en [Dashtbozorg et al., 2015], se utiliza un filtro de banda deslizante (SBF, del inglés *Sliding Band Filter*), una variante de los denominados filtros de índices de convergencia cuya salida mide el grado de convergencia de los vectores gradiente en una región dada. El filtro SBF se aplica en dos pasos. En primer lugar, se aplica a una versión de la imagen en baja resolución. La localización con la máxima respuesta será el centro aproximado del DO, a partir del cual se considera una región de interés (ROI). En segundo lugar, se aplica un SBF a la misma ROI, pero ahora en alta resolución para, de esta forma, afinar en el proceso de segmentación. La máxima respuesta del filtro da una estimación aproximada del contorno del DO. Finalmente se aplica un algoritmo de suavizado para obtener una segmentación más precisa de dicho contorno. De otro lado, en [Lu and Lim, 2010b], se aplica el denominado *filtro de Savitzky-Golay* [Savitzky and Golay, 1964] para extraer el background de la imagen (retina). El método se basa en la propiedad de que la intensidad de la retina es cuasi-constante mientras que el área del DO corresponde a una zona de alto brillo. Acto seguido, para obtener una aproximación del DO (detección), a la imagen original, se le resta la estimación del fondo. Esta primera aproximación corresponderá a un DO fragmentado en diferentes regiones, delimitadas por los vasos contenidos en el interior del DO. A fin de obtener una única región continua y de borde redondeado, la aproximación obtenida es transformada mediante operaciones morfológicas de erosión y dilatación.

Los algoritmos evolutivos se han usado también en la búsqueda del contorno papilar. Así, en [Carmona et al., 2008], se aplican técnicas de umbralización para detectar el DO y algoritmos genéticos para segmentarlo. En primer lugar, basándose en el alto brillo que presenta la zona papilar se obtiene el punto más brillante del canal rojo, detectando así la zona papilar. Dicho punto marca el centro de extracción de una subventana donde se realizará la segmentación. Posteriormente, en dicha subventana, se identifica la zona perteneciente al DO con mayor brillo (excavación) y, a partir del centro de la excavación, se trazan segmentos radiales dirigidos hacia la parte exterior de la papila. La idea es analizar estos segmentos para encontrar puntos, denominados puntos hipótesis (PH), en los que se produce un cambio brusco de intensidad (transición DO-retina). Finalmente, utilizando un algoritmo genético y una población de elipses (individuos), se realiza la búsqueda de aquella elipse que mejor se ajusta al contorno papilar. La función *fitness* devolverá un valor tanto más óptimo cuanto mayor sea el número de PHs contenidos en el contorno de la elipse. Siguiendo con el paradigma evolutivo, en [Novo et al., 2009], se utiliza lo que se denomina redes activas topológicas (TAN, del inglés *Topological Active Network*). Esta técnica aúna características de técnicas de segmentación basadas en regiones y bordes con el uso de modelos deformables, los cuales se ajustan a la topología del objeto que se quiere segmentar. Cada TAN se codifica como un individuo de la población en un algoritmo genético. La búsqueda evolutiva permite encontrar la TAN que mejor se adapta a las características del DO.

El uso combinado de la técnica *template matching* (detección) con otros métodos de segmentación ha dado también buenos resultados. Así, por ejemplo, en [Yu et al., 2012], se detecta el DO utilizando un *template* adaptado a distintas resoluciones. A partir de la posición aproximada del DO, se aplica la segmentación en una serie progresiva de pasos. En primer lugar, se realiza la extracción de una ROI en el canal rojo y centrada en el punto de detección, para posteriormente eliminar los vasos y las posibles lesiones brillantes. Acto seguido se aplica el denominado método de conjunto de niveles (LSM, del inglés *Level Set Method*), donde, a fin de adaptarse al contorno papilar, una curva cerrada es deformada basándose en la información local de bordes de imagen. Finalmente se realiza una estimación elíptica del contorno por mínimos cuadrados. Este método parece funcionar bien, incluso en

contornos borrosos de DO.

### Fóvea

Atendiendo al criterio establecido en la Sección 1.2.1, no se han podido encontrar referencias en la literatura en las que la segmentación de fóvea se haga usando solo CRintra-e. Esto puede estar motivado por el hecho de que tanto la fóvea como la mácula carecen de un contorno tan definido como el DO o los vasos y, por ello, la mayoría de los trabajos dedicados a la segmentación de fóvea toman como referencia la detección o segmentación de otra estructura previa (DO o red de vasos) para, posteriormente, hacer uso del CRinter-e a la hora de abordar la segmentación de esta estructura.

### Red de vasos

La red de vasos es una estructura compleja de segmentar, debido a que su nivel de gris no siempre es constante a lo largo de dicha estructura. Se han usado muchas técnicas para su segmentación [Ricci and Perfetti, 2007, Nekovei and Sun, 1995, Marin et al., 2011, Soares et al., 2006, McInerney and Terzopoulos, 1995, Alonso-Montes et al., 2008, Nyul, 2009, Aquino, 2014, Hoover et al., 2000, Giachetti et al., 2011, Hooshyar and Khayati, 2010, Calvo et al., 2011] y, prácticamente, todas ellas utilizan solo CRintra-e.

Por ejemplo, en muchos casos, el problema de segmentación de la red vascular se ha abordado como un problema de aprendizaje supervisado, en el que cada píxel se clasifica como perteneciente o no a vaso. Así, por ejemplo, el clasificador ha sido obtenido mediante diferentes métodos: máquinas de vectores soporte (SVM, del inglés *Support Vector Machines*) [Ricci and Perfetti, 2007], redes neuronales [Nekovei and Sun, 1995, Marin et al., 2011] o clasificadores Bayesianos [Soares et al., 2006]. Las aproximaciones basadas en contornos activos también han resultado útiles [McInerney and Terzopoulos, 1995, Alonso-Montes et al., 2008]. En otros trabajos, se utiliza la información geométrica y conocimiento morfológico de la red de vasos [Nyul, 2009, Aquino, 2014]. También se han usado técnicas de umbralización a partir de la respuesta de determinados filtros de imagen: *matched filter* [Hoover et al., 2000]

o *top-hat* [Giachetti et al., 2011]. En otros casos, se han utilizado métodos combinados, tal y como ocurre en [Hooshyar and Khayati, 2010], donde se combina el método de optimización por colonia de hormigas (ACO, del inglés *Ant Colony Optimization*) y el denominado agrupamiento difuso (del inglés, *fuzzy clustering*). Igualmente, en [Calvo et al., 2011], también se combinan distintas técnicas como los filtros de imagen, operadores morfológicos y un clasificador de puntos de vaso.

A modo de resumen, todos los métodos comentados a lo largo de esta sección solo utilizan CRintra-e, es decir, ninguno de ellos hace uso del CRinter-e. Desde este punto de vista, resulta difícil evaluar automáticamente si el resultado de segmentación corresponde a un verdadero positivo o a un falso positivo, dado que la estructura segmentada no utiliza nunca el contexto del resto de estructuras de la retina como marco de referencia.

### 2.3.2. Uso conjunto de CRintra e inter-e

Debido a la valiosa información que proporciona el CRintra-e, los trabajos que segmentan una o varias estructuras suelen utilizar conjuntamente ambos tipos de conocimiento relacional: intra e inter-e. Entre ellos, podemos establecer dos categorías: aquellos que segmentan más de una estructura de forma secuencial y aquellos otros que lo hacen de forma simultánea. La inmensa mayoría de los trabajos de segmentación de fovea hacen uso del CRinter-e, apoyándose normalmente en el DO y en la red vascular; de otro lado, el DO suele segmentarse utilizando solo CRintra-e, aunque, en contadas excepciones, también lo hace utilizando el CRinter-e.

#### Segmentación secuencial

La segmentación secuencial consiste en segmentar/detectar una o más estructuras previas para, a partir de estas, segmentar una o varias estructuras objetivo. En este apartado se citan los pocos trabajos que segmentan el DO a partir de otras estructuras (normalmente a partir del HV) y aquellos otros, más numerosos, que segmentan la fovea, apoyándose normalmente en la segmentación previa del DO y/o la red de vasos.

### *Segmentación secuencial dirigida al disco óptico*

No existen muchos trabajos que utilicen conjuntamente CR<sub>intra</sub> e inter-e para segmentar el DO. Como ya se ha comentado en la sección 2.3.1, esta estructura suele ser detectada en una primera etapa y, luego, en una segunda etapa, segmentada. Sin embargo, al contrario que ocurre con la segmentación de fóvea, la segmentación de DO no suele utilizar como referencia la posición de otra estructura. No obstante, la segmentación del DO utilizando como referencia el centro del HV [Welfer et al., 2010, Salazar-Gonzalez et al., 2014], previamente localizado, es una de las pocas excepciones que rompen con la tendencia mencionada. Hay que tener en cuenta que el punto de unión de las arcadas superior e inferior (el centro del HV) suele encontrarse muy cerca del centro del DO. En la sección 2.2.2 se detalló la obtención del HV para las dos referencias mencionadas anteriormente. En particular, en [Welfer et al., 2010], a partir del HV previamente obtenido, se realiza la segmentación del DO aplicando la transformación *Watershed*. En [Salazar-Gonzalez et al., 2014], el HV se utiliza para delimitar una ROI. Seguidamente, la segmentación del DO se realiza utilizando dos métodos diferentes. En primer lugar, el método de reconstrucción de imagen MRF (del inglés *Markov Random Field*) elimina los vasos de la región del DO y realiza una reconstrucción de la misma para finalmente segmentarlo mediante *graph-cut*. El problema del método de reconstrucción de imagen MRF es el gran coste computacional que supone (de ahí que en este trabajo se aplique solo a una ROI). En segundo lugar, el DO es segmentado mediante el denominado método del factor de compensación (del inglés, *compensation factor method*), el cual utiliza el conocimiento de intensidad local de los vasos. La baja intensidad de los vasos que están dentro del DO provoca que el método *graph-cut* tienda a etiquetarlos como píxeles no pertenecientes al DO; por tanto, el método de compensación modifica la forma de etiquetar los píxeles de la red de vasos por parte del *graph-cut*. Es decir, se está compensando el comportamiento de segmentación de los píxeles de vasos al segmentar el DO. Finalmente, en [Marin et al., 2015], se detecta el DO a partir de la red de vasos y varios procesos de umbralización. Una vez detectada la región de píxeles candidatos de DO, se aplica un detector de bordes de Prewitt. Adicionalmente, haciendo uso del resultado anterior, del tamaño típico del radio de DO y de

la transformada circular de Hough, se obtiene el borde del DO.

### *Segmentación secuencial dirigida a la fovea*

La mayoría de los trabajos de segmentación de fovea utilizan ambos tipos de conocimiento (CRintra e inter-e), utilizando como referencia la posición del DO y, en otros casos, adicional o alternativamente, la red de vasos. Aquí, tal y como se estableció en la sección 1.2.1, solo se considerarán aquellos trabajos de segmentación de fovea que evalúan sus resultados para un umbral de distancia, medida desde el centro real de fovea y el obtenido, menor o igual a  $\frac{1}{2}R_{DO}$ .

En [Gegundez et al., 2013] se utiliza diferentes técnicas de filtrado, umbralización y detección de bordes para segmentar la fovea. En este trabajo, la detección de DO y la segmentación de los vasos retinianos se utilizan para segmentar el centro de la fovea, utilizando la conocida relación de distancia entre los centros de fovea y DO. Este trabajo asume como conocidas la posición del DO y la segmentación de la red de vasos. A partir de esta última, se realiza una estimación parabólica de dicha red, para así calcular el eje de simetría de las arcadas superior e inferior y que une el DO con la zona macular. Con toda esta información se extrae una ROI situada a la distancia típica DO-fovea. En esta subventana se realiza la búsqueda de la fovea. Básicamente, se aplica un proceso de detección de bordes con el objeto de calcular un mapa de contornos. Este mapa de contornos se puede interpretar como un mapa topográfico de niveles de gris, donde el nivel más bajo correspondería a la zona más oscura y, por tanto, al centro de la estructura que se busca. En el último paso, se descartan aquellas líneas de contorno que no son cerradas para, finalmente, obtener el centroide de las líneas de contorno que sí lo son. Dicho centroide corresponderá al centro de fovea.

En [Aquino, 2014] se utilizan criterios basados en rasgos anatómicos y visuales para segmentar la fovea, junto con técnicas de procesamiento morfológico. Como en el trabajo descrito anteriormente, aquí también se parte de que son previamente conocidos la posición y segmentación del DO y de los vasos, respectivamente. En primer lugar, se obtiene una estimación aproximada del centro de la fovea utilizando las propiedades geométricas conocidas con respecto al DO y a la red de vasos. A partir del DO y la red de vasos se calcula una región de

interés (ROI) situada sobre el probable eje DO-fóvea. En segundo lugar, se emplea un tipo de procesamiento morfológico para intentar mejorar la estimación del centro de la fóvea.

La técnica de *template matching* también se ha usado en la segmentación de fóvea [Kao et al., 2014, Yu et al., 2011]. Así, en [Kao et al., 2014], se utiliza el CRinter-e existente entre el DO, los vasos y la fóvea. En primer lugar, y a partir de un *template* de papila, se aplica *template matching* para obtener el centro aproximado del DO. Basándose en la idea de que la fóvea es una zona libre de vasos, se utiliza la información de los mismos para establecer el eje entre ambas estructuras. A continuación, y volviendo a aplicar el CRinter-e, se busca una zona de interés (ROI) situada a la distancia típica DO-fóvea y trasladada sobre el eje que une estas dos estructuras. Finalmente, se vuelve a aplicar el *template matching* sobre la ROI detectada para segmentar la fóvea. Igualmente, en [Yu et al., 2011], también se utiliza la información de DO para segmentar la fóvea. Dado que a priori no se conoce si el ojo es de tipo izquierdo o derecho, se extraen dos subventanas de lado  $4R_{DO}$ , situadas a la distancia típica DO-fóvea, una a izquierda y otra a derecha del DO. La posición exacta de la fóvea, será el punto de respuesta más baja a la salida del *template matching* de entre estas dos subventanas.

Todos los trabajos de segmentación descritos que realizan una segmentación basada en la detección o segmentación de una o más estructuras previas tienen la ventaja de reducir la posibilidad de obtener un falso positivo como solución, ya que usan el conocimiento relacional existente entre las diferentes estructuras. Sin embargo, también tienen el inconveniente de que un fallo en la detección o segmentación de al menos una de las estructuras usadas como referencia provocará con alta probabilidad un fallo en la segmentación de la estructura que se quiere segmentar. Tal y como ocurría en la tarea de detección, para minimizar este último inconveniente, el conocimiento relacional también debería usarse en el marco de una segmentación simultánea de más de una estructura. De esta forma, al realizar la búsqueda de forma simultánea, el conocimiento relacional no presupone la existencia de una estructura detectada o segmentada de antemano. Nótese que, en caso contrario, una segmentación basada en la búsqueda sucesiva de diferentes estructuras, no podrá usar el conocimiento relacional en la segmentación de la primera estructura, solo podrá usar el conocimiento relacional a partir de

la segunda estructura buscada, pero nunca en la primera.

### Segmentación simultánea

La segmentación simultánea consiste en segmentar dos o más estructuras al mismo tiempo. A pesar de las ventajas mencionadas anteriormente, existen muy pocos trabajos en la literatura que realicen este tipo de segmentación [Girard et al., 2016, Dashtbozorg et al., 2016, Tan et al., 2017]. Así, en [Girard et al., 2016], se realizan las tareas de detección del DO y la fovea de forma simultánea, para posteriormente segmentar el borde del DO. En primer lugar se preprocesa la imagen, eliminando los vasos y realizando una ecualización del histograma a fin de mejorar el contraste. El proceso de selección de candidatos de fovea consiste en marcar una serie de semillas (puntos) de forma uniforme a lo largo de la imagen. Estas semillas evolucionan iterativamente (se van moviendo) hacia el mínimo local de su vecindario. El proceso se repite hasta que no exista movimiento en las semillas, escogiendo dichos puntos como puntos candidatos de fovea. Por otra parte, para escoger los candidatos de DO, se procede de la misma forma que con la fovea, pero en este caso las semillas son círculos de distinto radio. Estas semillas evolucionan iterativamente hacia el máximo local de su vecindario. Por otra parte, se calcula una serie de mapas de imagen según distintos criterios: detección de vasos mediante un filtro de Canny, entropía, varianza y dos mapas de simetría radial, uno positivo y otro negativo. Una vez que se tienen los conjuntos candidatos de DO y fovea, a partir de los mapas de imagen se les asigna una puntuación. Finalmente se aplica una función de puntuación por parejas donde la distancia típica fovea-DO es también utilizada. La pareja con mejor puntuación determinará las posiciones aproximadas de ambas estructuras. Para segmentar el DO, en primer lugar se traza un círculo centrado en el DO y se transforma a coordenadas polares, posteriormente se aplica un clasificador K-medias para detectar los bordes en el espacio polar. Por último se realizan las siguientes acciones: (1) todos los puntos seleccionados por el clasificador se transforman a espacio cartesiano; (2) a partir de los puntos transformados, se generan mini-segmentos; y (3) la curva poligonal de segmentos se transforma en una curva suave que es la que proporciona el límite papilar.

En el apartado 2.2.2 se comentó un trabajo de detección simultánea de DO y fovea

[Dashtbozorg et al., 2016]. En este trabajo, la fovea se evalúa con un nivel de exactitud que cumple con nuestro criterio de segmentación de fovea. Por tanto, aunque respecto al DO solo se hace detección, se puede incluir aquí como una variante aproximada de la categoría de métodos que se abordan en esta subsección.

En [Tan et al., 2017] se utilizan redes neuronales convolucionales para realizar la segmentación simultánea. En este trabajo se hace uso de una única red neuronal convolucional de siete capas (en las cinco primeras, las neuronas se distribuyen espacialmente en 2D y, en las dos últimas, en 1D) para, de forma simultánea, segmentar el DO, la fovea, la red vascular y el fondo de imagen (retina). En el trabajo mencionado, la primera capa es la capa de entrada. A continuación, se dispone una secuencia de dos pares de capas en el que cada par se compone de una capa de convolución más una capa de muestreo. La última capa de muestreo se conecta a una sexta capa de neuronas ordenadas unidimensionalmente y ésta, finalmente, se conecta a la capa de salida de cuatro neuronas, clasificando cada píxel de la imagen en una de las cuatro clases mencionadas. Además, la capa de entrada está formada por tres canales de tamaño fijo ( $33 \times 33$ ) que permite introducir tres subventanas de imagen centradas en un píxel dado. Así, el canal  $i$ -ésimo escala una subventana  $n \times n$ , centrada en el píxel elegido, hasta la dimensión fija de  $33 \times 33$ , siendo  $n=7, 33$  y  $165$ , respectivamente, para  $i=1, 2$  y  $3$ . La idea subyacente en el uso de estos tres canales es que la red aprenda el entorno de cada píxel a distintas escalas. El problema de este método es el elevado tiempo de entrenamiento requerido (emplea 35 382 s por cada época de entrenamiento), siendo el tiempo necesario para la segmentación de las cuatro estructuras por cada imagen también muy elevado (10 min o 2035 s, dependiendo de si se ejecuta, respectivamente, usando o no GPUs). Sin embargo, como ventaja cabe decir que el CRintra e inter-e no necesita ser inyectado por el usuario, sino que es la propia red neuronal la que lo usa de forma automática durante la fase de entrenamiento para construir la red.

La Tabla 2.2 resume los trabajos de segmentación mencionados en esta sección permitiendo la categorización de los mismos en función de diferentes descriptores: método o métodos principales utilizados para la segmentación; tipo de estructuras segmentadas; carácter independiente, secuencial o simultáneo de la segmentación; estructura requerida como entrada

(obtenida mediante un método descrito en un paper diferente); tipo de conocimiento utilizado; y tiempo empleado. Los trabajos se presenta en la tabla ordenados cronológicamente, de más antiguo a más reciente. A partir de dicha tabla, se pueden realizar las siguientes observaciones: (1) la inmensa mayoría de trabajos que segmentan el DO se apoyan solo en el uso del CRintra-e, detectándolo en primera instancia para luego ser segmentado; (2) con los vasos ocurre algo similar, donde solo el CRintra-e es el único tipo de conocimiento utilizado para su segmentación; (3) las tendencias mencionadas en (1) y (2) empieza a cambiar en trabajos más recientes, donde se tiende a utilizar ambos tipos de conocimiento; (4) en relación a la fóvea, la mayoría de los trabajos utilizan tanto el CRintra-e como el CRinter-e para realizar la segmentación de esta estructura; (5) la tendencia más reciente es la de segmentar varias estructuras de forma secuencial haciendo uso de ambos tipos de conocimiento y donde, normalmente, la fóvea es la última estructura en ser segmentada; y (6) al igual que ocurría con la detección, la segmentación simultánea de más de una estructura ha sido escasamente utilizada en la literatura.

Existe también en la literatura diferentes trabajos dedicados a recopilar el estado del arte asociado a la segmentación de las distintas estructuras retinianas. Algunos de estos trabajos, puesto que abarcan detección y segmentación, ya fueron indicados al final de la sección 2.2: DO [Besenczi et al., 2016, Mary et al., 2016], fóvea [Besenczi et al., 2016] y red de vasos [Besenczi et al., 2016]. Por otra parte, existen otros trabajos que se centran en hacer una revisión de métodos dedicados únicamente a la segmentación del DO [Allam et al., 2015, Almazroa et al., 2015]. Sin embargo, en ninguna de las referencias mencionadas se analiza la tarea de segmentación desde el enfoque descrito aquí, es decir, desde el punto de vista del tipo de conocimiento usado (CRintra o inter-e) o de la naturaleza del proceso de detección realizado (una o más estructuras y, en este último caso, si el proceso de detección es secuencial o simultáneo).

Todas las referencias mencionadas a lo largo de esta capítulo son solo una muestra representativa de la cantidad sustancial de enfoques existentes en la literatura y que están dedicados tanto a la detección como a la segmentación de estructuras anatómicas retinianas. Ambos problemas son considerados por la comunidad científica como problemas abiertos,

### 2.3. SEGMENTACIÓN DE ESTRUCTURAS ANATÓMICAS RETINIANAS

---

dirigiéndose los esfuerzos hoy día hacia la obtención de métodos de detección y segmentación cada vez más rápidos y robustos.

Tabla 2.2: Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras segmentadas	Tipo segmentación	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Lee and Brady, 1991]	Contornos Activos ( <i>Snakes</i> )	DO	Independ.	Detección DO	CRintra-e	-
[Nekoei and Sun, 1995]	Redes Neuronales	Vasos	Independ.	-	CRintra-e	540
[McInerney and Terzopoulos, 1995]	Contornos Activos ( <i>Snakes</i> )	Vasos	Independ.	-	CRintra-e	-
[Morris and Donnison, 1999]	Contornos Activos ( <i>Snakes</i> )	DO	Independ.	Detección DO	CRintra-e	[10, 20]
[Mendels et al., 1999]	Contornos Activos ( <i>Snakes</i> )	Detección DO→DO	Independ.	-	CRintra-e	-
[Hoover et al., 2000]	Filtro <i>Matched Filter</i> y técnicas de umbralización	Vasos	Independ.	-	CRintra-e	300
[Osareh et al., 2002a]	Contornos Activos ( <i>Snakes</i> )	Detección DO→DO	Independ.	-	CRintra-e	-
[Chrástek et al., 2002]	Transformada circular de Hough y filtros de bordes	Detección DO→DO	Independ.	-	CRintra-e	-
[Tang et al., 2003]	<i>Image Fusion</i>	Detección DO→DO	Independ.	-	CRintra-e	-
[Lowell et al., 2004]	Contornos Activos ( <i>Snakes</i> )	Detección DO→DO	Independ.	-	CRintra-e	-
[Tang et al., 2006]	Contornos Activos basados en regiones (Chan&Vese)	DO	Independ.	Detección DO	CRintra-e	-
[Soares et al., 2006]	Clasificadores Bayesianos	Vasos	Independ.	-	CRintra-e	190

Tabla 2.2: Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores (Cont.)

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras segmentadas	Tipo segmentación	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Ricci and Perfetti, 2007]	SVM ( <i>Support Vector Machine</i> )	Vasos	Independ.	-	CRintra-e	-
[Abdel-Ghafar and Morris, 2007]	Transformada circular de Hough	DO	Independ.	-	CRintra-e	-
[Kim et al., 2007]	Estimador matemático RANSAC ( <i>RANdom SAMple Consensus</i> ) y umbralización binaria	Detección DO→DO	Independ.	-	CRintra-e	-
[Carmona et al., 2008]	Algoritmos Genéticos	Detección DO→DO	Independ.	-	CRintra-e	[16.9, 80.8]
[Nyu, 2009]	Información geométrica y conocimiento morfológico de la red de vasos	Vasos	Independ.	-	CRintra-e	-
[Novo et al., 2009]	TAN (Topological Active Network)	Detección DO→DO	Independ.	-	CRintra-e	[20, 30]
[Joshi et al., 2010a]	Contornos Activos basados en regiones (Chan&Vese)	Detección DO→DO	Independ.	-	CRintra-e	-
[Lu and Lim, 2010b]	Filtro de Savitzky-Golay y operaciones morfológicas	Detección DO→DO	Independ.	-	CRintra-e	-
[Siddalingaswamy and Prabhu, 2010]	Contornos Activos ( <i>Snakes</i> )	Detección DO→DO	Independ.	-	CRintra-e	-
[Zhu et al., 2010]	Transformada circular de Hough	DO	Independ.	-	CRintra-e	30.26
[Hooshyar and Khayati, 2010]	ACO ( <i>Ant Colony Optimization</i> ) y agrupamiento difuso ( <i>Fuzzy Clustering</i> )	Vasos	Independ.	-	CRintra-e	-

Tabla 2.2: Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores (Cont.)

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras segmentadas	Tipo segmentación	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Aquino et al., 2010]	Transformada circular de Hough, operaciones morfológicas y filtros de bordes	Detección DO→DO	Independ.	-	CRintra-e	5.69 ± 0.54
[Welfer et al., 2010]	Conocimiento morfológico de la red de vasos y transformación <i>Watershed</i>	Vasos→HV→DO	Secuenc.	-	CRintra/inter-e	4.28
[Giachetti et al., 2011]	Filtro <i>Top-Hat</i> y técnicas de umbralización	Vasos	Independ.	-	CRintra-e	-
[Yu et al., 2011]	<i>Template Matching</i>	Detección DO→DO→fóvea	Secuenc.	-	CRintra/inter-e	9.8
[Yu et al., 2012]	<i>Template Matching</i> y LSM (Level Set Method)	Detección DO→DO	Independ.	-	CRintra-e	6.6
[Morales et al., 2013]	Análisis PCA ( <i>Principal Component Analysis</i> ), operaciones morfológicas y transformación <i>Watershed</i>	Detección DO→DO	Independ.	-	CRintra-e	-
[Gegundez et al., 2013]	Estimación parabólica y detectores de bordes	(Vasos, DO)→fóvea	Secuenc.	Detección DO y vasos	CRintra/inter-e	0.94 ± 0.54
[Aquino, 2014]	Morfología y criterios basados en rasgos anatómicos y visuales	(Vasos, DO)→fóvea	Secuenc.	Detección DO y vasos	CRintra/inter-e	10.88 ± 0.19
[Kao et al., 2014]	<i>Template Matching</i>	(Detección DO, vasos)→fóvea	Secuenc.	-	CRintra/inter-e	[2.6, 6.9]
[Salazar-Gonzalez et al., 2014]	MRF ( <i>Markov Random Field</i> ), <i>Compensation Factor Method</i> y <i>graph-cut</i>	Vasos→HV→DO	Secuenc.	-	CRintra/inter-e	-

Tabla 2.2: Categorización de diferentes trabajos de segmentación atendiendo a varios descriptores (Cont.)

Trabajo	Método(-s) y/o estrategia(-s) usadas	Estructuras segmentadas	Tipo segmentación	Entradas requeridas	Conocimiento usado	Tiempo procesado (s)
[Dashibozorg et al., 2015]	SBF ( <i>Sliding Band Filter</i> )	Detección DO→DO	Independ.	-	CRintra-e	[8.0, 11.3]
[Marin et al., 2015]	Transformada circular de Hough y filtro de bordes de Prewitt	Vasos → Detección DO→DO	Secuenc.	-	CRintra/inter-e	5.42 ± 0.98
[Dashibozorg et al., 2016]	SBF ( <i>Super Elliptical Filter</i> )	DO y fovea	Simult.	-	CRintra/inter-e	2.4
[Girard et al., 2016]	Niveles de gris, filtros de bordes, y clasificador K-medias	DO y fovea	Simult.	-	CRintra/inter-e	6.4
[Tan et al., 2017]	Red Neuronal Convolucional	DO, fovea y vasos	Simult.	-	CRintra/inter-e	2035

# Capítulo 3

## Metodología

En este capítulo se expone y desarrolla la metodología de detección y segmentación de estructuras anatómicas propuesta en esta tesis. Una característica común e importante de ambas metodologías es que el proceso de búsqueda de las estructuras detectadas o segmentadas se realiza siempre de forma simultánea y, además, utilizando siempre ambos tipos de conocimiento: CRintra e inter-e. El capítulo queda estructurado de la siguiente forma. En primer lugar, en la sección 3.1, se detallan las diferentes etapas de la metodología de detección: (i) preprocesado; (ii) obtención de blobs-estructuras mediante el uso de CRintra-e; y (iii) creación de tuplas-estructuras y obtención de la mejor tupla mediante el uso de CRinter-e. Seguidamente, en la sección 3.2, se detallan las diferentes etapas asociadas a la metodología de segmentación: (i) preprocesado; (ii) obtención de modelos intra-e; (iii) obtención de un modelo inter-e; y (iv) uso de un algoritmo evolutivo que, haciendo uso del modelo inter-e, permite segmentar simultáneamente varias estructuras. Paralelamente a la descripción de ambas metodologías, se describe también la implementación de sendos métodos, uno de detección y otro de segmentación, cada uno de ellos basado en su respectiva metodología. Finalmente, con el objetivo de mejorar el coste computacional y/o los resultados de segmentación, en la sección 3.3, se describe cómo modificar el método de segmentación mencionado anteriormente para aceptar y procesar el resultado obtenido por el método de detección.

### 3.1. Detección de estructuras retinianas

El objetivo de la metodología aquí propuesta es obtener la detección simultánea de las principales estructuras anatómicas en imágenes digitales de retina: DO, mácula, red de vasos y HV (ver Fig. 1.1). La idea básica es utilizar tanto el CRintra-e como el CRinter-e. La Fig. 3.1 muestra el esquema seguido. En primer lugar, se aplica una etapa de preprocesado en la que la imagen original es normalizada y reducida de tamaño. Luego, por cada tipo de estructura anatómica, se aplica un operador, el cual utiliza solo CRintra-e, a fin de obtener un conjunto de blobs candidatos, siendo cada uno de ellos una potencial solución al tipo de estructura buscado. Se utiliza el término *conjunto de candidatos*, porque cada conjunto debería contener cero o más falsos positivos y, también, el verdadero positivo buscado. Los falsos positivos estarían asociados a la existencia de distractores en la imagen, esto es, patrones similares a la estructura detectada. Por otra parte, el verdadero positivo corresponderá a la estructura solución que estamos buscando.

Una vez obtenidos los diferentes conjuntos de candidatos, se crea un conjunto de tuplas, donde cada tupla contiene una combinación de los diferentes tipos de blobs candidatos. A continuación, se aplican dos nuevos operadores, los cuales utilizan CRinter-e. El objetivo del primer operador es chequear diferentes relaciones asociadas al CRinter-e para detectar blobs falsos positivos a fin de eliminar todas aquellas tuplas que los contengan. Finalmente, la misión del segundo operador es seleccionar la mejor tupla, donde cada componente contendrá respectivamente el blob solución de cada estructura anatómica.

A continuación, se describe con más detalle cada una de las etapas implicadas en la metodología de detección aquí propuesta.

#### 3.1.1. Preprocesado

Los dos objetivos de esta etapa de preprocesado son normalizar y reducir el tamaño de la imagen de entrada. La normalización de intensidad se implementa mediante un método de expansión de contraste (*stretching*), es decir, la idea es que cada canal RGB de la imagen de entrada cubra el rango máximo de intensidad en el intervalo  $[0,255]$ . Por otra parte, la

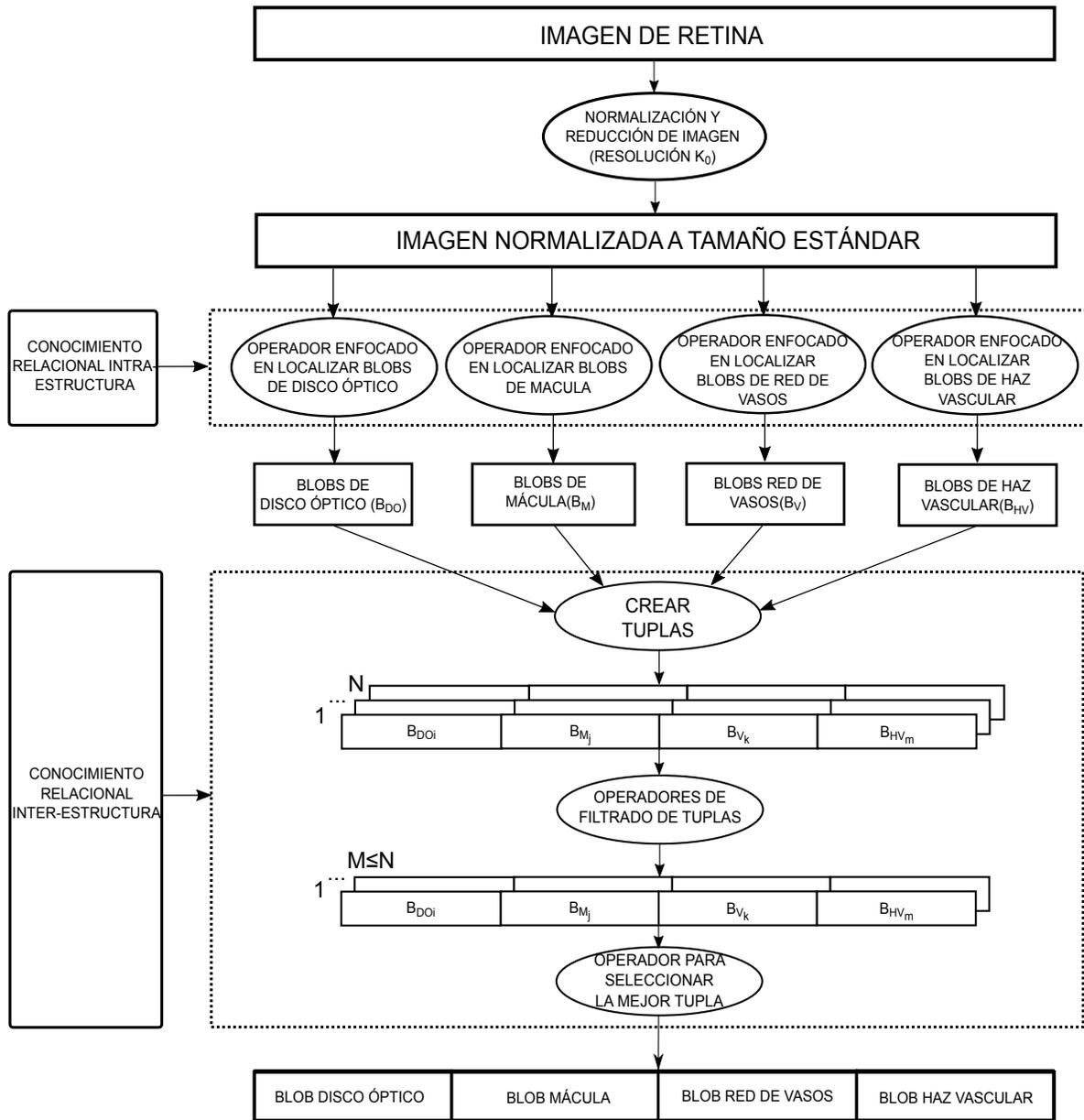


Figura 3.1: Diagrama de bloques de la metodología utilizada en detección.

### 3.1. DETECCIÓN DE ESTRUCTURAS RETINIANAS

---

reducción de tamaño se consigue utilizando un factor de escala, el cual está orientado a obtener un tamaño de resolución fijo de imagen. Este proceso de reducción tiene varias ventajas notables. En primer lugar, ayuda a decrementar el tiempo de procesamiento por imagen y, además, elimina posibles artefactos ruidosos. En segundo lugar, el uso de una resolución fija permitirá implementar y configurar un método de detección que siempre trabaje a esta resolución, es decir, un método independiente del tamaño original y de la resolución de la imagen de entrada. Esto es una gran ventaja debido a la gran variabilidad de tamaños y resoluciones de las bases de imágenes. Así, asumiendo que el método trabajará con una resolución estándar  $K_0$ , definida por el usuario y expresada en  $pixel/mm$ , es fácil establecer la siguiente relación:

$$\alpha = K_0 \sqrt{\frac{A_{mm^2}^\theta}{A_{px}}} \quad (3.1)$$

donde  $A_{mm^2}^\theta$  y  $A_{px}$  representan el área de la imagen de entrada, expresada en  $mm^2$  y píxeles, respectivamente, y  $\alpha$  es el factor de escala necesario para transformar la imagen de entrada a la resolución elegida ( $K_0$ ).  $A_{mm^2}^\theta$  depende del campo de visión (FOV) de la imagen de entrada, el cual también se conoce como el *ángulo de visión externo de la cámara de fondo de ojo*. La Tabla 3.1 muestra diferentes valores de  $A_{mm^2}^\theta$  para valores típicos de FOV [Atkinson and Mazo, 2011].  $A_{px}$  se puede calcular fácilmente a partir de la imagen de entrada. Para ello, bastaría con contar el número de píxeles reales de la retina. Finalmente, a partir de la Ec. (3.1), conocidos  $K_0$ ,  $A_{mm^2}^\theta$  y  $A_{px}$  es directo obtener el valor de  $\alpha$ . Por ejemplo, asumiendo que una imagen de retina fue tomada con una cámara cuyo FOV es  $40^\circ$ , entonces, consultando la Tabla 3.1,  $A_{mm^2}^\theta = 99.2 mm^2$ . Si el diámetro del círculo que rodea la zona retinal de la imagen es de 400 píxeles, tenemos que  $A_{px} = \pi(400/2)^2 = 125\,663.7 px^2$ . Finalmente, si el método de detección es configurado para, por ejemplo,  $K_0 = 10 px/mm$ , entonces, a partir de la Ec. (3.1), obtenemos que el factor de escala necesario para reducir el tamaño de la imagen original es de  $\alpha = 0.281$ .

Tabla 3.1: Área retinal para diferentes valores de ángulos del campo de visión (FOV), de acuerdo con [Atkinson and Mazo, 2011].

FOV	Área retinal ( $mm^2$ )
30°	56.4
40°	99.2
45°	124.8
50°	153.1

### 3.1.2. Utilizando conocimiento relacional intra-estructura

El CR<sub>intra-e</sub> hace referencia a toda la información que nos permite definir una estructura anatómica específica como, por ejemplo, su forma, color, textura o contorno. Aquí, la idea es usar diferentes operadores CR<sub>intra-e</sub>, uno para cada estructura a detectar. El resultado final de esta etapa es un conjunto de blobs candidatos para cada estructura. En cuanto a cómo implementar los diferentes operadores, se propone el uso de dos técnicas bien conocidas en el campo de la visión artificial: el *template matching* (TM) y el procesamiento morfológico de imágenes. Sin embargo, la metodología presentada aquí es independiente de las dos técnicas mencionadas, es decir, se podría haber utilizado cualquier otro enfoque orientado a obtener un conjunto de blobs candidatos para cada estructura. El resto de esta subsección describe cómo se implementan estos operadores intra-e en función de la estructura a detectar.

#### Detección de blobs del disco óptico

Se propone realizar la detección del conjunto de blobs candidatos de DO mediante la comentada técnica de TM con correlación. Así, una vez que la imagen de entrada es normalizada y reducida, se le aplica el TM. Para cada canal, el resultado de este proceso es una imagen de niveles de gris con la propiedad de tener valores altos en aquellas áreas más similares al *template* buscado. Finalmente, por cada canal se extraen los  $n_{do}$  blobs más brillantes, obteniendo un total de  $N_{do} = 3 \times n_{do}$  blobs candidatos de DO, donde  $n_{do}$  es elegido

### 3.1. DETECCIÓN DE ESTRUCTURAS RETINIANAS

---

por el usuario y el número tres corresponde a cada uno de los canales de imagen RGB. En el proceso de selección de blobs, se tiene en cuenta la propiedad de que un blob es una región continua aislada. El proceso de extracción de blobs comienza eligiendo un umbral igual a la intensidad del píxel más brillante de la salida del TM. Este píxel es denominado *blob actual*. A continuación, se va disminuyendo progresivamente ese umbral y, en cada decremento, se verifica si los nuevos píxeles con intensidad mayor o igual que el umbral actual son contiguos o no al *blob actual*. En el primer caso, los nuevos píxeles se incorporan al *blob actual* y la iteración continúa. En el segundo caso, el *blob actual* se guarda como *blob candidato*, el *blob actual* se asocia a la nueva área aislada y la iteración continúa. Este proceso se repite hasta que se obtienen  $n_{do}$  blobs candidatos o hasta que el umbral alcanza un valor por debajo del valor medio del nivel de gris de la imagen de salida del TM. Los tres *templates* fueron elegidos con una forma cuadrada y con una anchura  $W_{DO} = 4R_{DO}$ , donde  $R_{DO}$  es el radio medio del DO, en píxeles, a resolución estándar ( $K_0$ ). Para un valor de  $K_0$  conocido, resulta fácil calcular el valor de  $R_{DO}$  a partir de la expresión (3.2), sabiendo que el radio medio del DO, expresado en *mm*, es  $R_{DO,mm} = 0.925 \text{ mm}$  [Hubbard et al., 1999].

$$R_{DO} = K_0 \cdot R_{DO,mm} \quad (3.2)$$

El *template* para cada canal fue calculado como la intensidad media de un conjunto de diferentes subventanas cuadradas, con anchura  $W_{DO}$ , extraídas de la base de imágenes MESSIDOR, a resolución  $K_0$ , y centradas en el centro geométrico del DO. En concreto, 100 imágenes no patológicas de ojo derecho fueron extraídas de forma aleatoria de MESSIDOR. Este número fue seleccionado experimentalmente (valores mayores de 100 fueron probados, pero no hubo mejora sustancial en el proceso de detección). Además, cada una de las imágenes anteriores se reflejó horizontalmente para generar otras 100 imágenes de ojo izquierdo. Así, la matriz *template* final fue creada a partir de un total de 200 imágenes. Los tres *templates* generados para cada uno de los canales RGB se muestran en la figura 3.2.

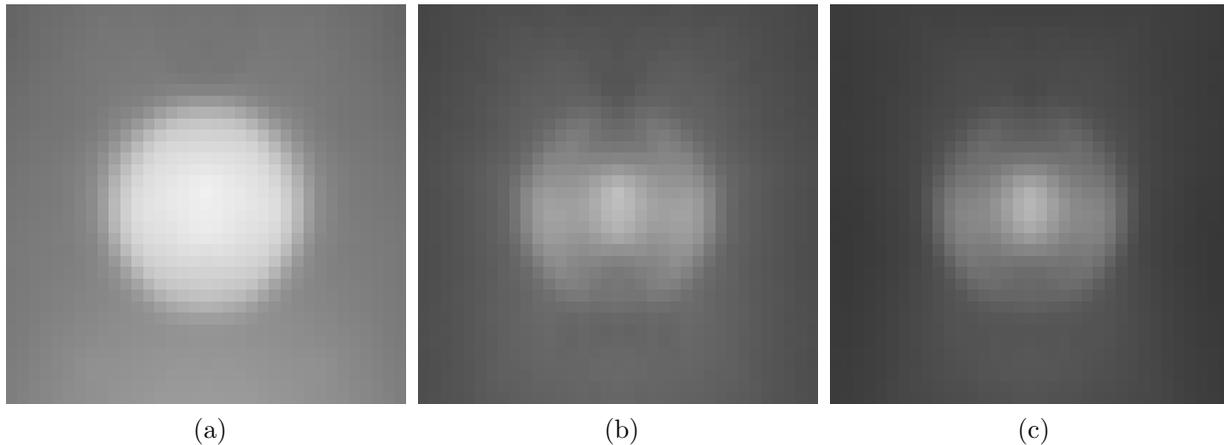


Figura 3.2: *Templates* de DO utilizados para cada canal RGB : (a) rojo, (b) verde, (c) azul.

### Detección de blobs de la mácula

Para detectar los blobs candidatos de mácula, se propone también utilizar el método TM con correlación. Como ya se hizo con el DO, se escogen los  $n_m$  blobs más brillantes por canal RGB de la salida del TM, creando un total de  $N_m = 3 \times n_m$  blobs candidatos de mácula, donde  $n_m$  es escogido por el usuario. En este caso, debido a que el área cubierta por la mácula es más pequeña que la del DO, el ancho elegido del *template* fue de  $W_M = 2R_{DO}$ , donde  $R_{DO}$  es obtenido de la Ec. (3.2). La matriz *template* para cada canal fue calculada como la intensidad media de un conjunto de diferentes subventanas cuadradas, con anchura  $W_M$ , extraídas de la base de imágenes MESSIDOR, a resolución  $K_0$ , y centradas en el centro real de la mácula. Se utilizaron las mismas 200 imágenes empleadas para la creación de los *templates* de DO. Los tres *templates* finales utilizados se muestran en la Fig. 3.3.

### Detección de la red de vasos principales

En una imagen de retina, la red de vasos sanguíneos tiene una geometría en forma de árbol (ver Fig. 1.1). Esta red puede ser dividida en dos partes: arcada superior e inferior. Ambas convergen hacia un punto localizado muy cerca del centro del DO, el centro del HV. Hay diferencias morfológicas entre los vasos principales y secundarios. De una parte, los primeros son más anchos, rectos y verticales que los segundos, de forma que los vasos principales, a

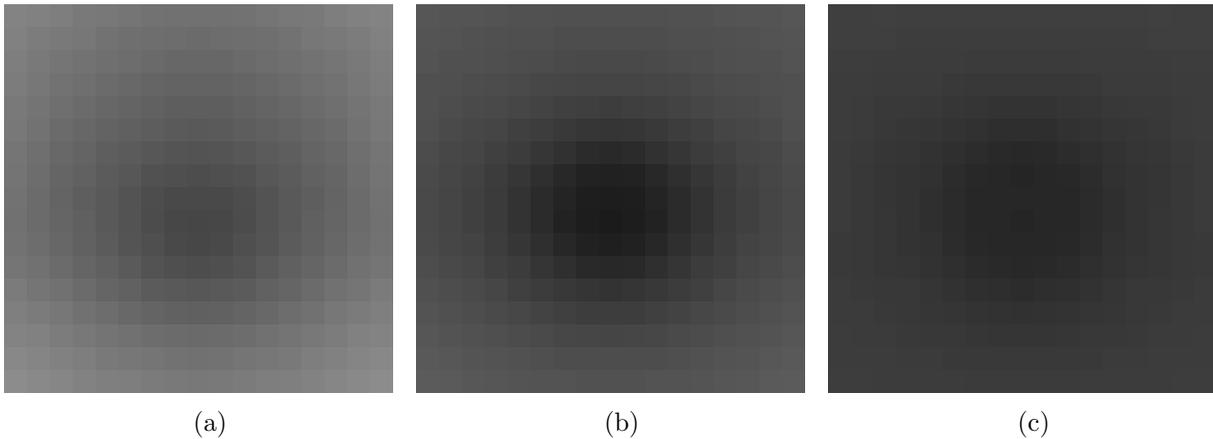


Figura 3.3: *Templates* de mácula utilizados para cada canal RGB : (a) rojo, (b) verde, (c) azul.

medida que se van alejando del centro del HV, decrecen en anchura, aumentan en ramificación y su orientación es más horizontal que vertical, convirtiéndose así en vasos secundarios. Aquí, a fin de realizar una localización aproximada de la red de vasos, solo estamos interesados en la detección de los vasos principales.

El proceso de detección que proponemos está basado en tres etapas morfológicas. Primero, se aplica una *transformación de escala de gris tipo bottom-hat*. En segundo lugar, se aplica una etapa de binarización con umbral iterativo. Finalmente, se aplica una etapa de posprocesado a fin de refinar la red de vasos retinales obtenida en el paso anterior.

La transformación *bottom-hat* es una operación morfológica que consiste en dos pasos: primero, a una imagen de entrada se le aplica una clausura morfológica; segundo, la imagen de entrada es sustraída del resultado de la clausura. En nuestro caso, la clausura se aplica únicamente al canal verde de  $I_0$  (imagen de entrada una vez normalizada y reducida), y el resultado,  $I_c$ , es una nueva imagen de retina sin vasos. Utilizamos solamente el canal verde porque existen evidencias experimentales en la literatura sobre el alto contraste existente entre vasos y el fondo para este canal [Lee et al., 2010]. La diferencia de la clausura y la imagen original de entrada,  $I_1 = I_c - I_0$ , produce una imagen en nivel de gris donde solo los vasos son realzados. El *elemento estructurante* utilizado para la clausura tiene forma circular

y su radio es calculado como la media aproximada del máximo diámetro de los vasos,  $\bar{d}_{vmax}$ , obtenido de la base de imágenes DRIVE [Staal et al., 2004] y considerando la resolución  $K_0$ . Utilizamos DRIVE porque el *ground truth* de la red de vasos está disponible públicamente en la web para esta base de imágenes [DRIVE, 2004]. El valor calculado para  $\bar{d}_{vmax}$  a la resolución original de DRIVE (sin aplicar el factor de escala  $\alpha$ ) fue de 9 px. La Fig. 3.4(a-c) muestra un ejemplo de los diferentes pasos de esta etapa.

Posteriormente, se toma  $I_1$  como entrada a la etapa de binarización con umbral iterativo. En este proceso, un umbral es decrementado progresivamente,  $th_{i+1} = th_i - \varepsilon$ , donde  $\varepsilon \gtrsim 0$  y  $th_0 = \max(I_1)$ . En cada iteración, el umbral  $th_i$  es utilizado para binarizar  $I_1$ , obteniendo de esta forma una aproximación binaria a la red de vasos. A medida que  $th_i$  se va haciendo cada vez más pequeño, el número de vasos mostrados en la imagen binaria es mayor. El proceso iterativo finaliza cuando el valor del área de vasos  $A_v$  es  $\bar{A}_v^{Drive} < A_v < 1.1\bar{A}_v^{Drive}$ , donde  $\bar{A}_v^{Drive}$  es el área media de la red de vasos para todas las imágenes de la base de imágenes DRIVE a la resolución  $K_0$ . El valor calculado para  $\bar{A}_v^{Drive}$  al tamaño original de DRIVE fue de 28 689 píxeles. La Fig. 3.4(d-f) muestra un ejemplo de los diferentes pasos de esta etapa.

Finalmente, se aplica una etapa de posprocesado a la salida de la segunda etapa,  $I_2$ . En concreto, se aplican cuatro operadores de forma secuencial. Primero, se aplica un filtrado para eliminar posibles blobs pequeños y ruidosos, utilizando un umbral de área como criterio de filtrado ( $th_{p1} = 0.001\bar{A}_v^{Drive}$ ). En segundo lugar, se aplica una clausura binaria para interconectar blobs de vasos aislados. En tercer lugar, se realiza un filtrado a fin de eliminar blobs sin aspecto de vaso. En concreto, este criterio de filtrado de blobs se basa en el cumplimiento simultáneo de tres propiedades relacionadas con las propiedades geométricas de un blob no-vaso: (i) la relación entre el área de blob y la del rectángulo que lo circunscribe es mayor que un umbral ( $th_{p2} = 0.2$ ); (ii) la relación entre el área del blob y el área cuadrada de lado  $\Delta x_{bb}$  es mayor que un umbral ( $th_{p3} = 0.2$ ), donde  $\Delta x_{bb}$  es la anchura de la caja que delimita el blob; y (iii) el área del blob es menor que un umbral ( $th_{p3} = 0.5\bar{A}_v^{Drive}$ ). La primera propiedad se utiliza para filtrar blobs ruidosos asociados con exudados y manchas algodonosas; la segunda, para filtrar manchas horizontales alargadas (vasos secundarios); y, la tercera, para filtrar blobs medianos o pequeños. Utilizando el conocimiento del dominio, se

escogió un valor inicial para cada parámetro  $th_{p1}$ ,  $th_{p2}$  y  $th_{p3}$ . A continuación, se realizaron diferentes experimentos para seleccionar los mejores valores de un conjunto de valores cercanos a los valores iniciales. El paso final de esta etapa de posprocesado consiste en aplicar una operación AND binaria, entre la última salida del operador de filtrado e  $I_2$ , con el fin de eliminar el efecto de dilatación de la clausura, pero preservando al mismo tiempo los efectos de filtrado. La Fig. 3.4(g-i) muestra un ejemplo de los diferentes pasos de esta etapa. A modo de resumen, el diagrama de bloques de la Fig. 3.5 muestra el proceso completo, donde se incluyen las tres etapas mencionadas arriba.

#### Detección de blobs del haz vascular

En una imagen de retina, el HV está formado por los vasos que están localizados en el interior del DO [de la Fuente-Arriaga et al., 2014]. El punto que enlaza las arcadas superior (nasal y temporal) e inferior (nasal y temporal), las cuales emergen del DO, es el llamado *centro del haz vascular*. Este punto está localizado muy cerca del centro del DO (ver Fig.1.1).

Aquí se propone realizar también la detección de los blobs candidatos del HV mediante TM con correlación. Sin embargo, a diferencia del DO y la mácula, aquí el TM no se aplica a las imágenes de retina RGB. En este caso, se aplica a imágenes de redes de vasos obtenidas con el método descrito en la sección anterior. Además, dado que la orientación de los vasos que salen del HV es diferente si trabajamos con el ojo izquierdo o con el derecho, se hace necesario crear un *template* de HV para cada tipo de ojo. Así, se eligen los  $n_{hv}$  blobs más brillantes en la salida del TM, obteniendo un total de  $N_{hv} = 2 \times n_{hv}$  blobs candidatos de HV, donde  $n_{hv}$  es elegido por el usuario y el número dos hace referencia a los *templates* de ojo izquierdo y derecho aplicados.

Los dos *templates* fueron elegidos con una forma cuadrada. Debido al hecho de que el *ground truth* del centro del HV no está disponible, primero, se calculó una matriz *template* auxiliar para ojo derecho como el promedio de intensidad de diferentes subventanas cuadradas, con anchura  $W' = 11R_{DO}$ , donde  $R_{DO}$  es obtenido de la Ec. (3.2). Estas subventanas se extraen de las imágenes binarias de la red del vasos y se centran en el centro geométrico real del DO. Elegimos un valor de  $W'$  lo suficientemente grande como para definir una subventana

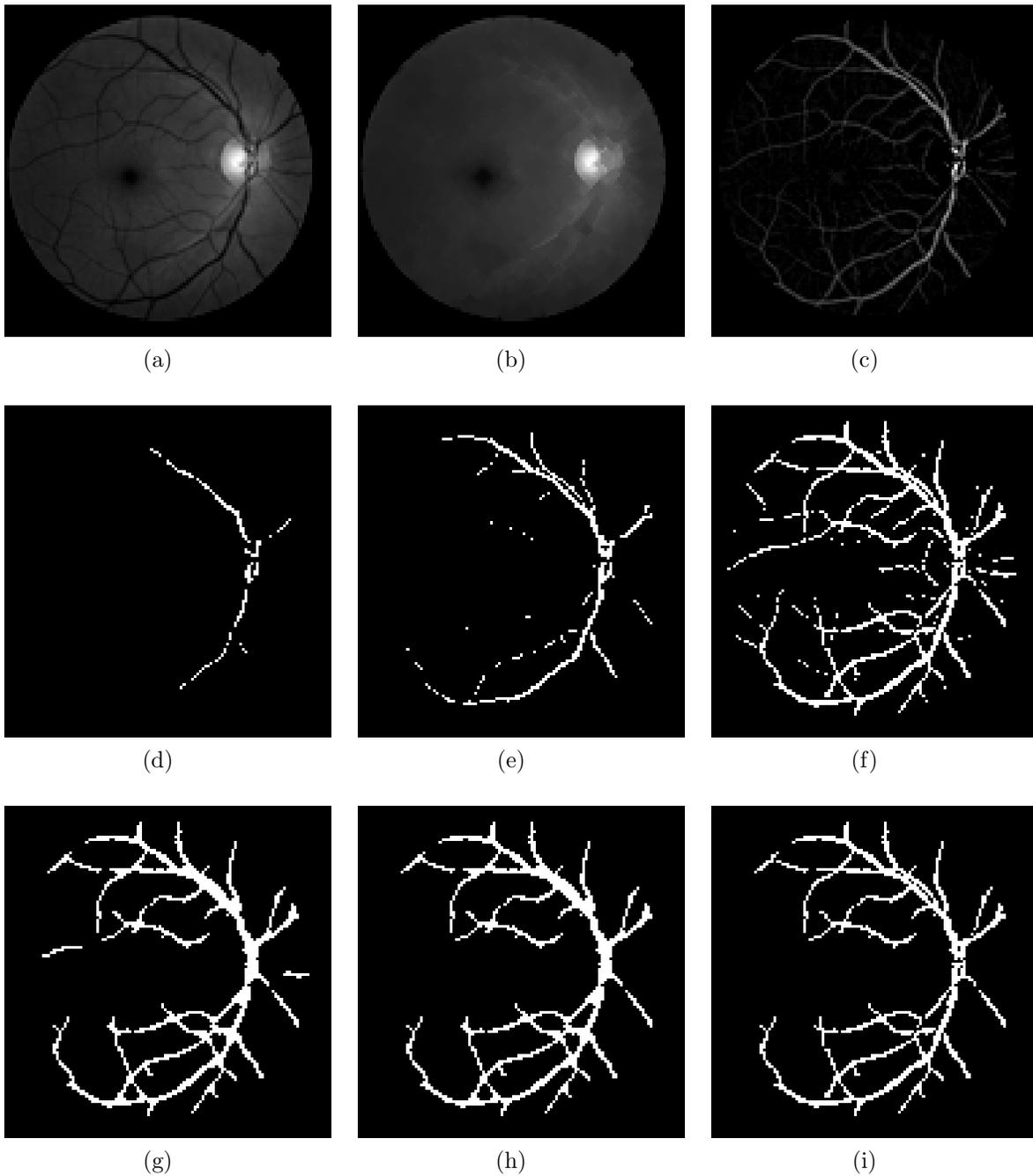


Figura 3.4: Ejemplo de detección de la red de vasos principales utilizando tres etapas: transformación *bottom-hat* (1ª fila), binarización de umbral iterativo (2ª fila) y etapa de posprocesamiento (3ª fila): (a)  $I_0$ , canal verde de la imagen normalizada y reducida (imagen de entrada); (b)  $I_c$ , clausura morfológica de  $I_0$ ; (c)  $I_1 = I_c - I_0$ , red de vasos en nivel de gris (salida de la primera etapa); (d-e) resultados de aplicar un umbral iterativo,  $th_i$ , a  $I_1$ ; (f)  $I_2$ , primera aproximación a la red binaria de vasos (salida de la segunda etapa); (g) filtrado y clausura de  $I_2$ ; (h) filtrado de los blobs con aspecto distinto a un vaso; (i) deshaciendo la clausura (red de vasos final).

### 3.1. DETECCIÓN DE ESTRUCTURAS RETINIANAS

---

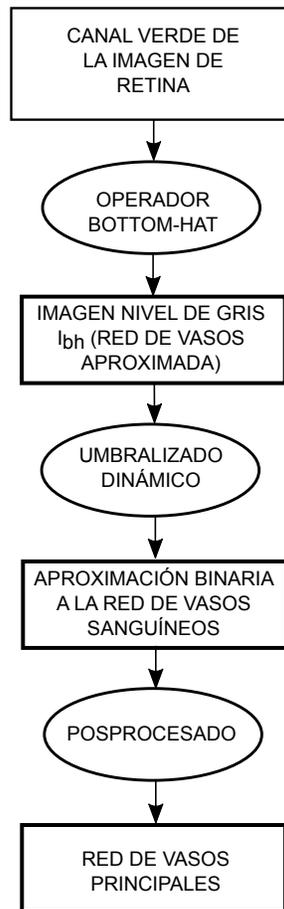


Figura 3.5: Diagrama de bloques que ilustra las tres etapas en el proceso de detección de la red de vasos principales.

que contenga y distinga visualmente la parte inicial de las cuatro arcadas que emergen del HV (superior, inferior, nasal y temporal). Dado que la localización real del centro del HV está normalmente localizada en una posición ligeramente desplazada del centro DO (ver, por ejemplo, la Fig. 1.1), el *template* final se obtiene extrayendo una subventana cuadrada del *template* auxiliar anterior y centrada visualmente en el centro del HV. Para lograr esto, el tamaño final de la nueva subventana,  $W_{HV}$ , tiene que ser ligeramente menor que  $W'$  (elegimos  $W_{HV} = 10R_{DO}$ ).

Las imágenes de la red de los vasos utilizadas aquí son el resultado de la aplicación del método de detección de vasos principales (véase la sección 3.1.2) a diferentes imágenes de MESSIDOR. En concreto, se utilizaron 200 imágenes no patológicas pertenecientes a ojo derecho. El *template* de ojo izquierdo se obtuvo reflejando horizontalmente el *template* obtenido de ojo derecho. La Fig. 3.6 muestra los dos *templates* de HV finalmente utilizados (ojo izquierdo y derecho).

Adicionalmente, el uso de los dos *templates* de HV tiene una ventaja adicional: puede utilizarse como un método sencillo y rápido para determinar si la imagen analizada contiene un ojo izquierdo o derecho. Para ello, bastará elegir el tipo de *template* (izquierdo o derecho) que provoca la máxima respuesta del TM para una imagen dada. De esta manera, el tipo de *template* ganador determinará el tipo de ojo.

Finalmente, a modo de resumen, la Fig. 3.7 muestra el diagrama de flujo de los diferentes operadores utilizados para obtener los conjuntos de blobs candidatos para las diferentes estructuras anatómicas.

### 3.1.3. Utilizando conocimiento relacional inter-estructura

Una de las principales contribuciones de esta metodología es el uso del CRinter-e a fin de eliminar blobs candidatos falsos positivos, y simultáneamente detectar las diferentes estructuras anatómicas (verdaderos positivos): DO, mácula y HV. Así, se utilizan las siguientes relaciones entre las principales estructuras anatómicas de retina:

- $R_1$ : El centro geométrico del DO está localizado dentro del área abarcada por él mismo

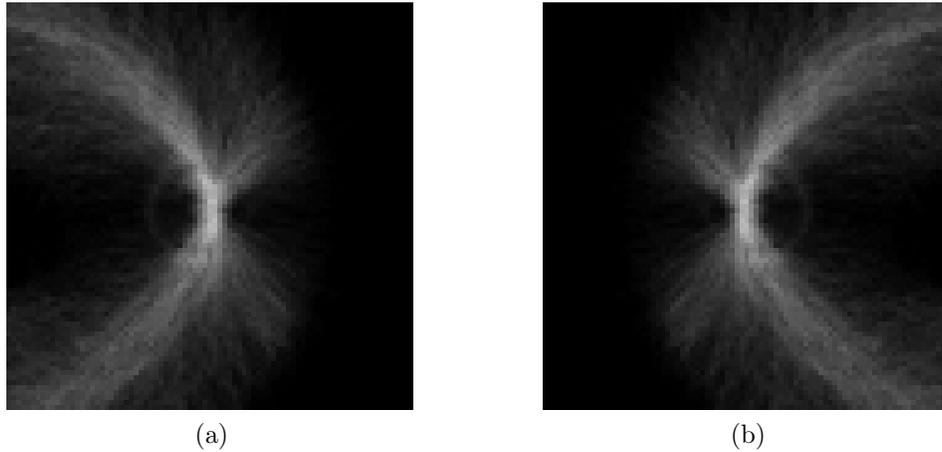


Figura 3.6: *Templates* de haz vascular utilizados para cada tipo de ojo: (a) ojo derecho; (b) ojo izquierdo.

y donde, además, siempre existe una alta acumulación de vasos (ver Fig. 1.1).

- $R_2$ : El centro de la mácula se encuentra en una región avascular, es decir, una región donde no hay vasos retinianos [Giachetti et al., 2013].
- $R_3$ : El centro del HV es un punto perteneciente a un vaso, rodeado de más vasos, localizado dentro del área del DO, y situado muy cerca del centro geométrico del DO (ver Fig. 1.1).
- $R_4$ : El centro de la mácula se encuentra generalmente a una distancia de aproximadamente 5 veces el  $R_{DO}$ , medido desde el centro del propio DO [Sinthanayothin et al., 1999]. Esta distancia es denotada como  $\tilde{D}(DO, F)$ :

$$\tilde{D}(OD, F) = 5R_{DO} \tag{3.3}$$

- $R_5$ : En el ojo derecho, el HV y el DO están siempre localizados a la derecha de la mácula (ver Fig. 1.1). En el ojo izquierdo ocurre justo al contrario.
- $R_6$ : En el ojo derecho, el centro del HV está normalmente localizado a la derecha del centro geométrico del DO (ver Fig. 1.1). En el ojo izquierdo ocurre justo al contrario.

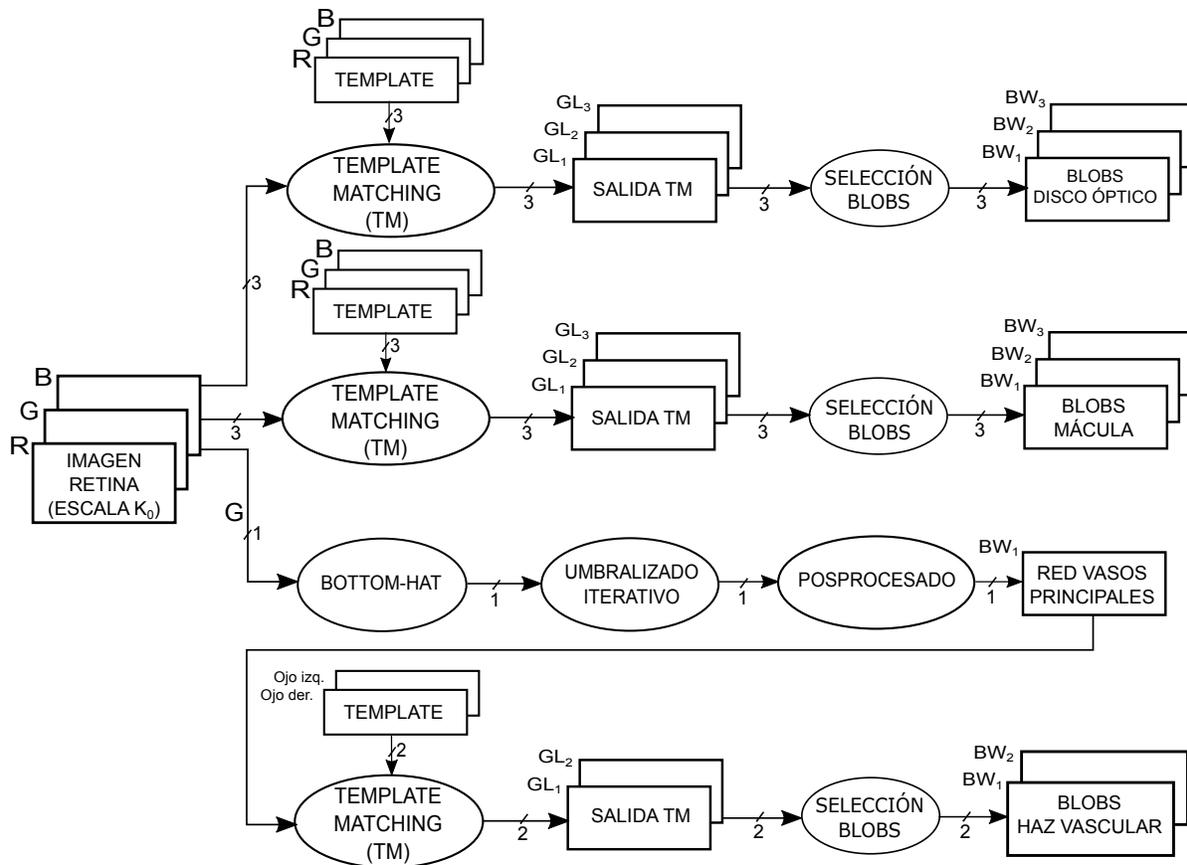


Figura 3.7: Diagrama de bloques que muestra las diferentes etapas utilizadas para obtener los diferentes conjuntos de blobs candidatos: DO, mácula, vasos y HV. Los canales de color están representados como  $R$  (rojo),  $G$  (verde) y  $B$  (azul); las imágenes de niveles de gris están representadas como  $GL_i$ ; y las imágenes en blanco y negro (imágenes binarias) están representadas como  $BW_i$ .

- $R_7$ : El ángulo que forma la horizontal con el segmento que une los centros de DO y mácula está acotado:  $(6.11 \pm 3.32)^\circ$  [Williams and Wilkinson, 1992],  $(5.6 \pm 3.3)^\circ$  [Rohrschneider, 2004] o  $(7.76 \pm 3.63)^\circ$  [Jonas et al., 2015].

El CRinter-e arriba mencionado será utilizado para seleccionar la mejor tripleta  $(B_{DO}, B_M, B_{HV})$ . Así, asumiendo que  $N_{DO}$ ,  $N_M$  y  $N_{HV}$  son, respectivamente, el número total de blobs candidatos de DO, M y HV, el algoritmo de selección de la mejor tripleta se implementa de la siguiente manera:

1. **Eliminación de blobs DO falsos positivos:** cada blob DO cuyo centro geométrico se encuentra a una distancia de cualquier vaso retiniano mayor que un valor umbral es descartado (ver relación  $R_1$ ). Tomamos  $1R_{DO}$  como valor umbral (ver Ec. (3.2)). A continuación,  $N_{DO}$  es actualizado adecuadamente.
2. **Descarte de blobs mácula falsos positivos:** cada blob de mácula cuyo centro geométrico se encuentra a una distancia de cualquier vaso retiniano menor que un valor umbral es descartado (ver relación  $R_2$ ). Tomamos  $\frac{1}{4}R_{DO}$  como valor umbral. A continuación,  $N_M$  es actualizado adecuadamente.
3. **Descarte de blobs HV falsos positivos:** cada blob HV cuyo centro geométrico se encuentra a una distancia de cualquier vaso retiniano mayor que un valor umbral es descartado (ver relación  $R_3$ ). Tomamos  $1R_{DO}$  como valor umbral. A continuación,  $N_{HV}$  es actualizado adecuadamente.
4. **Creación de tripletas:** con los blobs restantes, se crea un conjunto formado por todas las posibles tripletas,  $(B_{DO_i}, B_{M_j}, B_{HV_k})$ , donde  $i = 1, \dots, N_{DO}$ ,  $j = 1, \dots, N_M$  y  $k = 1, \dots, N_{HV}$ .
5. **Descarte de tripletas falsas positivas:** cada tripleta que contenga una pareja de blobs que verifique cualquiera de las siguientes condiciones es descartada del conjunto de tripletas.

- a)  $D(B_{DO_i}, B_{HV_k}) > 2R_{DO}$ , es decir, la distancia entre los centros geométricos de los blobs DO y HV es mayor que un valor umbral. Tomamos  $2R_{DO}$  como valor umbral porque es la distancia máxima permitida considerando que el DO y HV tienen que estar dentro del área de DO (ver las relaciones  $R_1$  y  $R_3$ ).
- b)  $\left| D(B_{DO_i}, B_{M_j}) - \tilde{D}(DO, F) \right| > 2R_{DO}$ , donde  $|\cdot|$  es el operador valor absoluto,  $D(B_{DO_i}, B_{M_j})$  es la distancia entre los centros geométricos de los blobs DO y mácula, y  $\tilde{D}(DO, F)$  es obtenido de la relación  $R_4$ . Tomamos  $2R_{DO}$  como valor umbral porque el peor de los casos correspondería a la situación en la que los blobs DO y mácula tuvieran una desviación de  $1R_{DO}$  en relación a sus respectivas posiciones reales y, además, se encontraran en posiciones opuestas.
- c)  $\left| D(B_{M_j}, B_{HV_k}) - \tilde{D}(DO, F) \right| > 2R_{DO}$ , donde  $D(B_{M_j}, B_{HV_k})$  es la distancia entre los centros geométricos de los blobs HV y mácula, y  $\tilde{D}(DO, F)$  es obtenido de la relación  $R_4$ . Esta condición es una consecuencia de la proximidad de los centros reales de DO y HV (ver relación  $R_3$ ). Por lo tanto, tomamos  $2R_{DO}$  como valor umbral por la misma razón explicada en el caso anterior.
- d) El blob mácula se encuentra a la izquierda del blob HV izquierdo o, a la inversa, el blob mácula se encuentra a la derecha del blob HV derecho. Ambas condiciones son consecuencia de  $R_5$ .
- e) El blob DO se encuentra a la izquierda del blob HV izquierdo y a una distancia mayor que  $1R_{DO}$  o, a la inversa, el blob DO se encuentra a la derecha del blob HV derecho y a una distancia mayor que  $1R_{DO}$ . Esta condición es una consecuencia de las relaciones  $R_3$  y  $R_6$ .
- f) El ángulo formado por la horizontal y el segmento que une los centros de mácula y DO está fuera del intervalo  $[-25, 25]^\circ$ . Se toma un intervalo lo suficientemente conservador para abarcar el peor caso contemplado en la relación  $R_7$ .
- g) El ángulo formado por la horizontal y el segmento que une los centros de mácula y HV está fuera del intervalo  $[-25, 25]^\circ$ . Esta condición es consecuencia de las relaciones  $R_3$  y  $R_7$ .

6. **Selección de la mejor tripleta:** si el conjunto final de tripletas obtenido en el paso 5 no está vacío, la siguiente función *fitness* es aplicada a cada tripleta superviviente del paso anterior:

$$F(B_{DO_i}, B_{M_j}, B_{HV_k}) = F_1(B_{DO_i}, B_{M_j}) + F_2(B_{M_j}, B_{HV_k}) + F_3(B_{DO_i}, B_{HV_k}) \quad (3.4)$$

con

$$F_1(B_{DO_i}, B_{M_j}) = \frac{I_{TM}(B_{DO_i}) \cdot I_{TM}(B_{M_j})}{1 + \frac{|D(B_{DO_i}, B_{M_j}) - \bar{D}(DO, F)|}{2R_{DO}}} \quad (3.5)$$

$$F_2(B_{M_j}, B_{HV_k}) = \frac{I_{TM}(B_{M_j}) \cdot I_{TM}(B_{HV_k})}{1 + \frac{|D(B_{M_j}, B_{HV_k}) - \bar{D}(DO, F)|}{2R_{DO}}} \quad (3.6)$$

$$F_3(B_{DO_i}, B_{HV_k}) = \frac{I_{TM}(B_{DO_i}) \cdot I_{TM}(B_{HV_k})}{1 + \left[ \frac{D(B_{DO_i}, B_{HV_k})}{2R_{DO}} \right]} \quad (3.7)$$

donde  $I_{TM}(B_{x_i})$  es la intensidad de nivel de gris de la salida TM para el blob  $i$ -ésimo de tipo- $x$ , y  $|\cdot|$  es el operador valor absoluto. Nótese que, en las Ec. (3.5), (3.6) y (3.7), los denominadores están normalizados por  $2R_{DO}$ , y los numeradores están escalados a  $[-1, +1]$ , debido al hecho de que la técnica TM que aquí se propone se basa en correlación cruzada normalizada. Entonces la tripleta que obtenga el máximo valor de la función *fitness* será la solución de detección a las tres estructuras anatómicas buscadas:

$$(B_{DO}, B_M, B_{HV}) = \arg \max_{x \in \mathbf{B}} F(x) \quad (3.8)$$

donde  $\mathbf{B} = \{(B_{DO_i}, B_{M_j}, B_{HV_k}), \forall i, j, k\}$ .

7. **Si el conjunto de tripletas queda vacío:** en otro caso (si el conjunto de tripletas está vacío), regresamos al paso 4, pero ahora considerando todas las parejas posibles,  $(B_{DO_i}, B_{M_j})$ ,  $(B_{M_j}, B_{HV_k})$  y  $(B_{DO_i}, B_{HV_k})$ . Acto seguido, las parejas falsas positivas son descartadas de acuerdo con las reglas expresadas en el paso 5. Una vez realizado el descarte, cada tipo de pareja se evalúa, respectivamente, de acuerdo con las funciones *fitness* (3.5), (3.6) y (3.7). Después de esto, la mejor pareja se selecciona de acuerdo al

criterio de máximo valor de función *fitness* para cada tipo de pareja:

$$(B_{DO}^1, B_M^1) = \arg \max_{x \in \mathbf{B}_1} F_1(x)$$

$$(B_M^2, B_{HV}^2) = \arg \max_{x \in \mathbf{B}_2} F_2(x)$$

$$(B_{DO}^3, B_{HV}^3) = \arg \max_{x \in \mathbf{B}_3} F_3(x)$$

donde:

$$\mathbf{B}_1 = \{(B_{DO_i}, B_{M_j}), \forall i, j\}$$

$$\mathbf{B}_2 = \{(B_{M_j}, B_{HV_k}), \forall j, k\}$$

$$\mathbf{B}_3 = \{(B_{DO_i}, B_{HV_k}), \forall i, k\}$$

En este paso, cada tipo de blob puede aparecer en dos de las tres parejas ganadoras o, en el peor de los casos,  $\mathbf{B}_1$ ,  $\mathbf{B}_2$  y  $\mathbf{B}_3$  pueden ser conjuntos vacíos. En todos estos casos, el blob solución para cada estructura se elige de acuerdo con las Ec. (3.9), (3.10) y (3.11).

$$B_{DO} = \begin{cases} \arg \max_{x=B_{DO}^1, y=B_{DO}^3} \{F_1(x, B_M^1), F_3(y, B_{HV}^3)\}, & \text{if } \mathbf{B}_1 \cup \mathbf{B}_3 \neq \emptyset \\ \arg \max_{x \in B_{DO_i}} I_{TM}(x) & \text{otro caso} \end{cases} \quad (3.9)$$

$$B_M = \begin{cases} \arg \max_{x=B_M^1, y=B_M^2} \{F_1(B_{DO}^1, x), F_2(y, B_{HV}^2)\} & \text{if } \mathbf{B}_1 \cup \mathbf{B}_2 \neq \emptyset \\ \arg \max_{x \in B_{M_j}} I_{TM}(x) & \text{otro caso} \end{cases} \quad (3.10)$$

$$B_{HV} = \begin{cases} \arg \max_{x=B_{HV}^2, y=B_{HV}^3} \{F_2(B_M^2, x), F_3(B_{DO}^3, y)\} & \text{if } \mathbf{B}_2 \cup \mathbf{B}_3 \neq \emptyset \\ \arg \max_{x \in B_{HV_k}} I_{TM}(x) & \text{otro caso} \end{cases} \quad (3.11)$$

### 3.2. Segmentación de estructuras retinianas

A continuación se describe la metodología propuesta para segmentar simultáneamente diferentes estructuras anatómicas, empezando por el caso de dos estructuras para, posteriormente, generalizar dicha metodología al caso de tres o más estructuras. Concretamente, la Fig. 3.8 muestra un diagrama de bloques que resume la operativa seguida en dicho método. En primer lugar, haciendo uso del CRintra-e asociado a la estructura #1, se genera un modelo llamado *modelo CRintra-e*. Se repite el proceso para la estructura #2, generando un segundo *modelo CRintra-e*. En ambos casos, se hace uso de una base de imágenes de retina en la que los *ground truth* de ambas estructuras son conocidos. Seguidamente, estos dos modelos CRintra-e se fusionan en un nuevo modelo denominado *modelo CRinter-e*, que incorpora el CRinter-e existente entre ambas estructuras. La idea final es transformar el problema de segmentar simultáneamente las estructuras #1 y #2 en un problema de optimización, consistente en encontrar los mejores parámetros de una instancia del modelo relacional que minimice el error respecto a dicho modelo. Debido al alto número de parámetros del modelo y al dominio de definición real de cada uno de ellos, no se puede usar un método de fuerza bruta para obtener la instancia óptima. Por lo tanto, se propone utilizar un algoritmo evolutivo para encontrar un ajuste óptimo (o casi óptimo). De esta forma, el mejor individuo de la población obtenido al final del proceso evolutivo representará la solución buscada. Hay que tener en cuenta que esta metodología podría usarse recursivamente en dos direcciones. Por ejemplo, por un lado, el modelo CRinter-e que resulta de combinar los dos modelos CRintra-e podría combinarse con el modelo CRintra-e de una tercera estructura para crear un nuevo modelo CRinter-e. Por otro lado, tres modelos CRintra-e podrían combinarse para crear un nuevo modelo CRinter-e. En ambos casos, el nuevo modelo CRinter-e nos permitiría

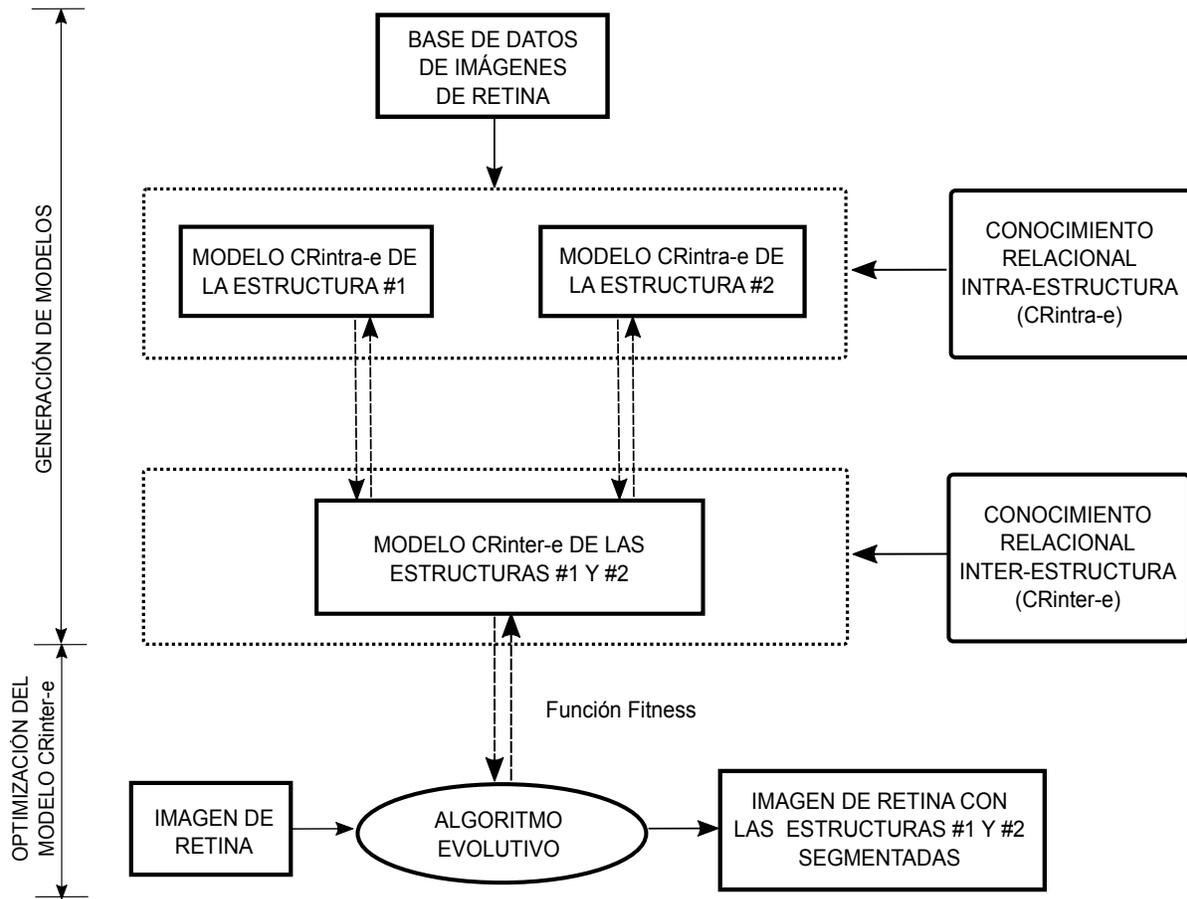


Figura 3.8: Diagrama de bloques asociado a la metodología usada para la segmentación simultánea de las estructuras #1 y #2.

segmentar simultáneamente tres estructuras, y así sucesivamente.

A través de esta sección, la metodología mencionada se utilizará para implementar un método dirigido a segmentar simultáneamente el DO y la fóvea. Los pasos seguidos para crear los diferentes modelos y las características del algoritmo evolutivo utilizado para resolver el problema de optimización se describirán en las siguientes subsecciones.

### 3.2.1. Preprocesado

En esta sección, se describen una serie de pasos comunes de preprocesamiento que serán utilizados tanto para la creación de cada uno de los modelos como para la evaluación de dichos modelos. Son los siguientes: (a) reducción del tamaño de la imagen de entrada para alcanzar

una resolución de trabajo común e independiente de la resolución de la imagen considerada; (b) transformación, si procede, de la imagen a un tipo de ojo elegido arbitrariamente por el usuario (izquierdo o derecho); (c) normalización de los canales RGB.

Para el paso de reducción de tamaño se sigue el mismo método que ya fue descrito en la sección 3.1.1, pero en este caso el valor elegido para la resolución de trabajo es  $K_0 = 25 \text{ px/mm}$ . Este valor no es crítico, pero debe ser lo suficientemente pequeño como para eliminar algunos artefactos ruidosos presentes a resoluciones mayores y, suficientemente grande, para que las estructuras a segmentar no pierdan el detalle necesario para ser segmentadas con éxito.

De otro lado, para hacer la segmentación de cada estructura retiniana independiente del tipo de ojo (izquierdo o derecho), se chequea, para cada imagen, el tipo de ojo y, si este no coincide con el tipo establecido por defecto, la imagen es reflejada horizontalmente. Se eligió el criterio, totalmente arbitrario, de trabajar siempre con ojo izquierdo. El método utilizado para determinar el tipo de ojo se describió ya en la sección 3.1.2. No obstante, se pueden utilizar otros métodos para abordar esta misma tarea [Tan et al., 2009, Roy et al., 2016].

Finalmente, para utilizar el máximo rango dinámico de cada canal RGB, se realizará una normalización en cada uno de ellos. Esto se consigue mediante un *contrast stretching method*, es decir, mediante una transformación lineal  $f : [min_{ret}, max_{ret}] \rightarrow [0, 255]$ , donde  $min_{ret}$  y  $max_{ret}$  corresponden a los valores de intensidad mínimo y máximo, respectivamente, del conjunto de píxeles de la imagen que solo forman parte de la retina, es decir, sin incluir el marco oscuro que normalmente presentan las imágenes de retina.

### 3.2.2. Modelo de Disco Óptico

Como ya se comentó en la sección 3.1.2, el DO es brillante y tiene forma ovalada, conteniendo además el punto de unión de los principales vasos retinianos, esto es, el centro del HV (ver Fig. 1.1). En ausencia de ruido, solo hay dos tipos de puntos de contorno papilar: puntos pertenecientes a la frontera retina-papila o puntos perteneciente a un vaso situado en la frontera retina-papila. De ahora en adelante, al primer tipo de puntos se denominará *punto frontera* y, al segundo, *punto vaso*. Sin embargo, dado que las imágenes de retina reales contendrán ruido, todos aquellos puntos del contorno de DO que no sean punto frontera o

vaso se denominarán *puntos ruido*.

A continuación, se propone la estimación de la probabilidad  $p$  de que un punto del contorno de DO, de acuerdo con su posición, pertenezca al conjunto de puntos frontera (la pertenencia al conjunto de puntos de vaso vendría dada entonces por  $1 - p$ ). Para la estimación de este conjunto de probabilidades, se siguen los pasos indicados en el algoritmo 3.1. Así, como información de entrada, se utiliza un conjunto de imágenes de retina,  $\mathbf{I}_i$ ,  $i = 1, \dots, N$ , donde el *ground truth* del contorno de DO ( $\mathbf{Cnt}_{DO_i}$ ) y del centro fóvea ( $\mathbf{C}_{F_i}$ ) son conocidos. Un `bucle-for` externo recorre todo el conjunto de imágenes y, para cada una de ellas, se realizan tres etapas bien diferenciadas: preprocesado, extracción de la red de vasos, y un conjunto de pasos encaminados a evaluar y almacenar la probabilidad de que cada punto del contorno papilar, dependiendo de su posición, sea un punto frontera (`bucle-for` interno). Finalmente, calculadas las probabilidades de los puntos del contorno papilar de todas y cada una de las imágenes, se calculan las probabilidades promedio.

La etapa de preprocesado sigue los pasos ya descritos en la sección 3.2.1 y en el siguiente orden: normalización de la imagen de entrada, transformación a una resolución común ( $K_0$ ) y, si procede, reflexión horizontal (incluido el *ground truth* de DO y fóvea) para unificar el tipo de ojo. Finalmente, la imagen RGB resultante, se transforma a niveles de gris,  $\mathbf{I}_{NG}$ . La segmentación aproximada de la red de vasos,  $\mathbf{V}$ , se realiza a partir del canal verde de la imagen RGB resultante del preprocesado y utilizando el método descrito en la sección 3.1.2. Solo utilizamos el canal verde porque hay evidencia experimental en la literatura sobre el alto contraste entre los vasos y el fondo para este canal [Lee et al., 2010].

Para facilitar la definición del modelo de DO, tal y como se verá al final de esta misma sección, se asume que el contorno de DO se puede aproximar mediante una elipse. Esta aproximación es común en la literatura [Carmona et al., 2008, Lowell et al., 2004, Pallawala et al., 2004, Cheng et al., 2011, Roychowdhury et al., 2016, Kusumandari et al., 2015, Ayub et al., 2016]. Por esta razón, tal y como se indica en la línea 5 del Algoritmo 3.1, el contorno real de DO proporcionado por el experto es aproximado por una elipse,  $e_{DO}$ , (ver Fig. 3.9.a). Además, la probabilidad de que cada punto del contorno papilar sea un punto frontera depende también de su posición. Por tanto, se hace necesario buscar un punto de partida en

---

**Algoritmo 3.1** Pseudocódigo para obtener el modelo de DO.

---

**Entradas**

- $\mathbf{I}_i$ , conjunto de imágenes ( $i = 1, \dots, N$ )
- $K_0$ , resolución de trabajo
- ojo*, tipo de ojo (izdo)
- $\mathbf{Cnt}_{DO_i}$ , *ground truth* de contorno de DO en imagen i-ésima
- $\mathbf{C}_{F_i}$ , *ground truth* del centro de fovea en imagen i-ésima
- $\delta$ , distancia entre elipse de DO ( $e_{DO}$ ) y elipse auxiliar ( $e_{aux}$ )
- $N_p$ ,  $N^0$  de puntos de muestreo en  $e_{DO}$  y  $e_{aux}$
- $n$ ,  $N^0$  de puntos vecinos a chequear ( $n + 1$ ) en  $e_{aux}$

**Salidas**

- $\mathbf{M}_{DO}$ , modelo de DO
- 

```

(01)  $\mathbf{t} := [N, \dots, N]_{1 \times N_p}$ ;
(02) for  $i := 1$  to  $N$ 
(03)    $[\mathbf{I}_{NG}, \mathbf{I}_G, \mathbf{Cnt}_{DO}, \mathbf{C}_F] := \text{Preprocesa}(\mathbf{I}_i, \mathbf{Cnt}_{DO_i}, \mathbf{C}_{F_i}, \text{ojo}, K_0)$ ;
(04)    $\mathbf{V} := \text{SegmentaVasos}(\mathbf{I}_G)$ ;
(05)    $e_{DO} := \text{AproximaElipse}(\mathbf{Cnt}_{DO})$ ;
(06)    $e_{aux} := \text{ConstruyeElipse}(e_{DO}, \delta)$ ;
(07)    $[P_{DO}, P_{aux}] := \text{CalculaPtosCorte}(e_{DO}, e_{aux}, \mathbf{C}_F)$ ;
(08)    $[\tilde{e}_{DO}, \tilde{e}_{aux}] := \text{MuestreaElipses}([e_{DO}, P_{DO}], [e_{aux}, P_{aux}], N_p)$ ;
(09)   for  $j := 1$  to  $N_p$ 
(10)     if  $\tilde{e}_{DO}(j) \in \text{Vaso}$ 
(11)        $\mathbf{p}(i, j) := 0$ ;
(12)     else if  $\tilde{e}_{DO}(j) \in \text{Frontera}$ 
(13)        $\mathbf{p}(i, j) := 1$ ;
(14)     else
(15)        $\mathbf{p}(i, j) := 0$ ;
(16)        $\mathbf{t}(j) := \mathbf{t}(j) - 1$ ;
(17)     end-if
(18)   end-for
(19) end-for
(20)  $\mathbf{M}_{DO} := [\sum_{i=1}^N (\mathbf{p}(i, 1)/\mathbf{t}(1)), \dots, \sum_{i=1}^N (\mathbf{p}(i, N_p)/\mathbf{t}(N_p))]$ ;

```

---

el contorno elíptico papilar para que el recorrido del mismo sea independiente del grado de inclinación del eje mayor de la elipse. Para ello (línea 7), se calculan los puntos de intersección de  $e_{DO}$  y  $e_{aux}$  con la recta que une los centros de DO y de fovea (ver Fig. 3.9.b). Estos puntos,  $P_{DO}$  y  $P_{aux}$ , marcarán el punto inicial del muestreo y almacenamiento de  $N_p$  puntos separados regularmente en  $e_{DO}$  y  $e_{aux}$ , respectivamente (línea 8).

De otro lado, es necesario operativizar el concepto de “pertenecer a un punto frontera”. Para ello, se chequean tres condiciones anidadas en una estructura `if-elseif-else` (líneas 10-17). Primero, se comprueba si dicho punto pertenece a un punto de vaso, utilizando para ello la imagen binaria de vasos segmentados  $\mathbf{V}$ . De ser así, se le asigna una probabilidad cero de pertenecer a un punto frontera. En otro caso, se chequea si el punto es un punto frontera. Para ello, se compara su intensidad con la de los  $n + 1$  puntos vecinos más cercanos y situados en una elipse auxiliar,  $e_{aux}$ , construida de forma concéntrica, exterior a la elipse-DO, y separada por un offset,  $\delta$  (ver Fig. 3.9.b). Entonces, si la intensidad del punto analizado es mayor que la de todos sus  $n + 1$  vecinos más cercanos en  $e_{aux}$ , se considerará un punto frontera y se le asignará una probabilidad de 1. Esta heurística se basa en la propiedad de que los puntos frontera del contorno papilar separan el interior del DO (zona brillante) de la zona retiniana (menos brillante) y, matemáticamente, se podría expresar así:

$$\tilde{\mathbf{e}}_{DO}(j) \in \text{Frontera} \iff \begin{cases} \mathbf{I}_{NG}(\tilde{\mathbf{e}}_{DO}(j)) > \mathbf{I}_{NG}(\tilde{\mathbf{e}}_{aux}(k)), \\ \forall k \in \{j - \frac{n}{2}, \dots, j, \dots, j + \frac{n}{2}\} \end{cases} \quad (3.12)$$

Finalmente, si el punto analizado no cumpliera ninguna de las dos condiciones anteriormente citadas, se trataría de un punto ruidoso y la contribución de este punto sería ignorada en el cálculo de la media final. Para ello, se le asignaría una probabilidad cero y, además, se restaría del cómputo total de puntos utilizados para calcular la probabilidad promedio (líneas 15-16). La Fig. 3.9.c muestra un ejemplo de etiquetado de los puntos de un contorno de DO que es aproximado por una elipse.

La base de imágenes MESSIDOR ( $N = 1200$ ) fue el conjunto utilizado para construir el modelo de probabilidades. El número de puntos de muestreo elegido fue  $N_p = 140$ , el valor

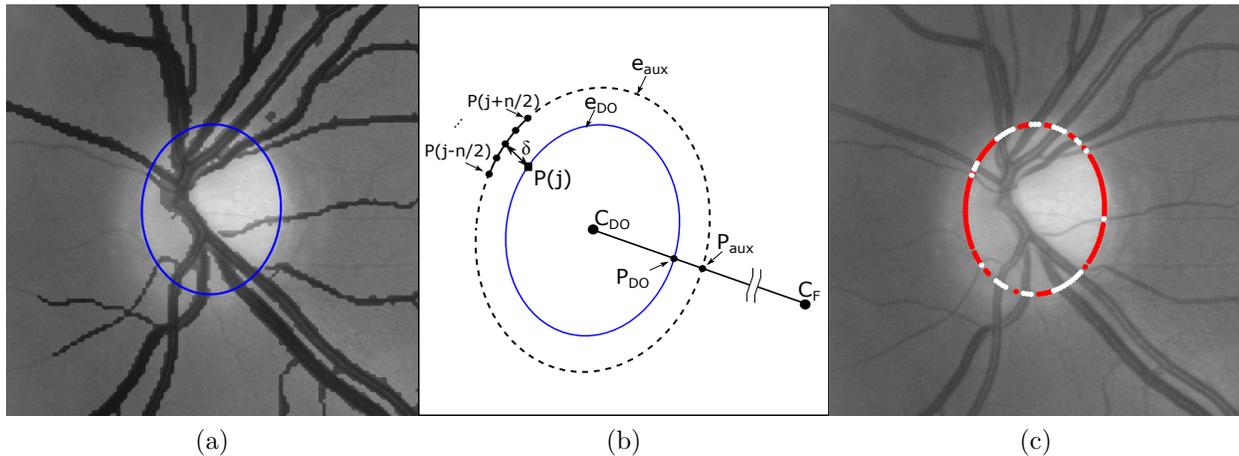


Figura 3.9: Ejemplo que muestra el etiquetado de los puntos que pertenecen al contorno de DO: (a) aproximando el contorno de DO mediante una elipse,  $e_{DO}$  (también se muestra la segmentación de vasos), (b) los  $n + 1$  puntos más cercanos, pertenecientes a  $e_{aux}$ , los cuales son utilizados para evaluar un punto de contorno,  $P(j)$ , en  $e_{DO}$  (los puntos de corte,  $P_{DO}$  y  $P_{aux}$ , también son mostrados), (c) ejemplo de clasificación de los puntos de contorno de DO como puntos frontera (rojo), puntos vaso (blanco) o puntos ruido (sin marca de color).

de separación entre  $e_{DO}$  y  $e_{aux}$  fue  $\delta = 2$  y el valor que determina el número de puntos vecinos ( $n + 1$ ) fue  $n = 8$ . Estos tres valores fueron obtenidos experimentalmente. No obstante, los valores de  $N_p$ ,  $\delta$  y  $n$  están relacionados con el valor de  $K_0$ , de tal forma que los primeros deben aumentar cuando el último aumenta y viceversa.

La Fig. 3.10.a muestra en trazo continuo la curva de probabilidades obtenida mediante el Algoritmo 3.1. El eje- $x$  representa los  $N_p$  puntos muestreados, empezando en  $P_{DO}$  y haciendo un recorrido de la elipse en sentido horario, y el eje- $y$  representa la probabilidad promedio de cada punto de pertenecer a la clase punto frontera. Se puede observar que los valores pico corresponden a la zona temporal y nasal, es decir, donde existe menor concentración de vasos y, por tanto,  $p \rightarrow 1$ . En cambio, los valores valle de la curva corresponden a la zona superior e inferior del DO ( $p \ll 1$ ), es decir, las zonas por donde la mayoría de los vasos abandonan el DO para, posteriormente, ramificarse por el resto de la retina. Finalmente, para filtrar el ruido de alta frecuencia de la curva de probabilidad obtenida, se realiza un suavizado de la misma mediante un desarrollo en serie de Fourier de diez armónicos. El resultado se muestra en la Fig. 3.10.b. De aquí en adelante, el vector de probabilidades que almacena secuencial

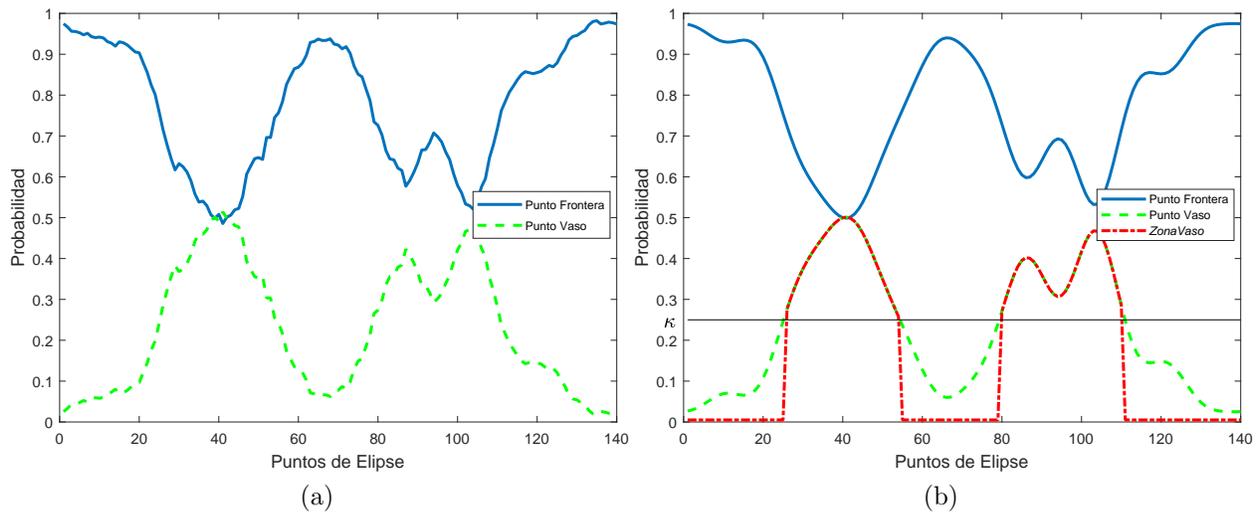


Figura 3.10: Modelo de DO ( $\mathbf{M}_{DO}$ ) representado por la curva de probabilidad promedio de que un punto de la elipse-DO sea un punto frontera: (a)  $\mathbf{M}_{DO}$  obtenido por el Algoritmo 3.1; (b)  $\mathbf{M}_{DO}$  final obtenido como resultado de aproximar la curva anterior con un desarrollo en serie de Fourier de 10 armónicos. En ambos casos, también se muestra la curva de probabilidad promedio de que un punto sea un punto vaso.

y ordenadamente las diferentes probabilidades de que un punto del contorno papilar sea un punto frontera se denominará *modelo de DO*. En la mencionada figura también se representa, en trazo discontinuo, la curva de probabilidades promedio  $(1-p)$  de que cada punto pertenezca a la clase punto de vaso.

### Cómo evaluar una instancia del modelo de DO

Obtenido el modelo de DO,  $\mathbf{M}_{DO}$ , la pregunta es cómo evaluar una instancia del modelo,  $\widehat{\mathbf{M}}_{DO}$ , es decir, cómo evaluar una elipse de parámetros  $(\widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho})$ , dada una imagen de retina,  $\mathbf{I}$ :

$$\widehat{\mathbf{M}}_{DO} = \widehat{f}_{DO}(\mathbf{I}, \widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho}, \widehat{\mathbf{C}}_F) \quad (3.13)$$

donde  $\widehat{\mathbf{C}}_F$  es una predicción del centro de fóvea y  $\widehat{f}_{DO}$  es la función de evaluación.

La implementación de  $\widehat{f}_{DO}$  viene dada por el Algoritmo 3.2, mostrando los pasos necesarios para evaluar  $\widehat{\mathbf{M}}_{DO}$ . En primer lugar, en la imagen de entrada, se realiza el mismo preprocesado

que el que se realizó para obtener  $\mathbf{M}_{DO}$ , manteniendo el mismo valor para los parámetros  $K_0$ ,  $ojo$ ,  $\delta$  y  $N_p$ . El resto del algoritmo corre casi paralelo al Algoritmo 3.1, con la diferencia de que las probabilidades evaluadas ahora pertenecen a los puntos de la elipse de entrada y se asignan de la siguiente forma. Si un punto  $j$  de la elipse es un punto frontera, es decir, cumple la condición expresada en (3.12), entonces se le asigna la probabilidad  $j$ -ésima del modelo. Si el punto es ruidoso, entonces se le asigna un valor cero para penalizar y, por tanto, maximizar el valor de error asociado a este tipo de puntos. Finalmente, si el punto  $j$ -ésimo es un punto de vaso, la asignación de probabilidad se realiza dependiendo del cumplimiento o no del siguiente criterio:

$$(\tilde{\mathbf{e}}_{DO}(j) \in ZonaVaso) \iff (1 - \mathbf{M}_{DO}(j)) \geq \kappa \quad (3.14)$$

donde  $(1 - \mathbf{M}_{DO}(j))$  representa, de acuerdo al modelo de DO, la probabilidad de que un punto sea punto de vaso, y  $\kappa \in [\min(1 - \mathbf{M}_{DO}(j)), \max(1 - \mathbf{M}_{DO}(j))]$  es un umbral fijo, elegido por el usuario, que establece la anchura de una zona en la parte superior e inferior del contorno de DO, que se denominará *ZonaVaso* (ver Fig. 3.10.b). Entonces, si un punto de vaso pertenece a *ZonaVaso*, se le asignará la probabilidad de vaso correspondiente (líneas 13-14). En otro caso, será considerado como un punto ruidoso y se le asignará una probabilidad cero (líneas 15-16). La finalidad de definir esta zona es penalizar elipses que en la zona temporal o nasal del contorno de DO siguen trayectorias pertenecientes a la frontera de un vaso. Cuanto mayor sea  $\kappa$ , más estrecha será *ZonaVaso* y, por tanto, menor será la posibilidad de elegir elipses con el problema mencionado. Sin embargo, si asignamos a  $\kappa$  el valor máximo de su dominio de definición, no podremos utilizar, en la búsqueda de la solución, la heurística de que la mayoría de los vasos cruzan el contorno de DO por su parte superior e inferior. Teniendo en cuenta ambos factores, los mejores resultados los hemos obtenido para  $\kappa = 0.4$ .

Finalmente, para evaluar cuánto se aproxima la instancia al modelo, bastará calcular:

$$error_{DO} = f_{errDO}(\mathbf{M}_{DO}, \widehat{\mathbf{M}}_{DO}) \quad (3.15)$$

donde  $f_{errDO}$  puede ser cualquier función de error, elegida por el usuario, que permita calcular

---

**Algoritmo 3.2** Pseudocódigo para evaluar una instancia del modelo de DO.

---

**Entradas**

- $\mathbf{I}$ , imagen de entrada
- $(\widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho})$ , valor de parámetros que definen la elipse a evaluar (instancia)
- $\widehat{\mathbf{C}}_F$ , valor del parámetro centro de fóvea
- $K_0, ojo, \delta, N_p, n$ , parámetros definidos en Algoritmo 3.1

**Salida**

- $\widehat{\mathbf{M}}_{DO}$ , evaluación de la instancia del modelo de DO
- 

```

(01)  $[\mathbf{I}_{NG}, \mathbf{I}_G] := \text{Preprocesa}(\mathbf{I}, \text{ojo}, K_0);$ 
(02)  $\mathbf{V} := \text{SegmentaVasos}(\mathbf{I}_G);$ 
(03)  $[e_{DO}, e_{aux}] := \text{ConstruyeElipses}([\widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho}], \delta);$ 
(04)  $[p_{DO}, p_{aux}] := \text{CalculaPtosCorte}(e_{DO}, e_{aux}, \widehat{\mathbf{C}}_F);$ 
(05)  $[\widetilde{\mathbf{e}}_{DO}, \widetilde{\mathbf{e}}_{aux}] := \text{MuestreaElipses}([e_{DO}, p_{DO}], [e_{aux}, p_{aux}], N_p);$ 
(06) for  $j := 1$  to  $N_p$ 
(07)   if  $(\widetilde{\mathbf{e}}_{DO}(j) \in \text{Frontera});$ 
(08)      $\widehat{\mathbf{M}}_{DO}(j) := \mathbf{M}_{DO}(j);$ 
(09)   end-if
(10)   if  $(\widetilde{\mathbf{e}}_{DO}(j) \notin \text{Vaso})$  AND  $(\widetilde{\mathbf{e}}_{DO}(j) \notin \text{Frontera})$ 
(11)      $\widehat{\mathbf{M}}_{DO}(j) := 0;$ 
(12)   end-if
(13)   if  $(\widetilde{\mathbf{e}}_{DO}(j) \in \text{Vaso})$  AND  $(\widetilde{\mathbf{e}}_{DO}(j) \in \text{ZonaVaso})$ 
(14)      $\widehat{\mathbf{M}}_{DO}(j) := 1 - \mathbf{M}_{DO}(j);$ 
(15)   else if  $(\widetilde{\mathbf{e}}_{DO}(j) \in \text{Vaso})$  AND  $(\widetilde{\mathbf{e}}_{DO}(j) \notin \text{ZonaVaso})$ 
(16)      $\widehat{\mathbf{M}}_{DO}(j) := 0;$ 
(17)   end-if
(18) end-if
(19) end-for

```

---

la discrepancia entre  $\mathbf{M}_{DO}$  y  $\widehat{\mathbf{M}}_{DO}$ .

### 3.2.3. Modelo de Fóvea

En una imagen de retina, los niveles de intensidad en la zona de fóvea crecen siguiendo la dirección de un haz imaginario situado en el punto central de la fóvea (ver Fig. 1.1). Aquí se propone caracterizar este patrón de variación de intensidades a partir de los ángulos y módulos de los vectores resultantes de aplicar el operador gradiente a una subventana centrada en fóvea.

Para la construcción de este patrón de fóvea, se siguen los pasos indicados en el Algoritmo 3.3. Así, como información de entrada, se utiliza un conjunto de imágenes de retina,  $\mathbf{I}_i$ ,

## 3.2. SEGMENTACIÓN DE ESTRUCTURAS RETINIANAS

---

---

**Algoritmo 3.3** Pseudocódigo para obtener el modelo de fovea.

---

**Entradas**

- $I_i$ , conjunto de imágenes ( $i = 1, \dots, N$ )
- $K_0$ , resolución de trabajo
- ojo*, tipo de ojo (izdo)
- $C_{F_i}$ , *ground truth* del centro de fovea en imagen  $i$ -ésima
- $L$ , tamaño de subventana de fovea ( $L \times L$ )

**Salidas**

- $M_F$ , modelo de fovea
- 

```
(01) for  $i:=1$  to  $N$ 
(02) [ $I_G, C_F$ ] := Preprocesa( $I_i, C_{F_i}, ojo, K_0$ );
(03)  $W_F$  := ExtraeSubventanaFovea( $I_G, C_F, L$ );
(04) [ $Gx(i), Gy(i)$ ] := CalculaGradiente( $W_F$ );
(05) end-for
(06)  $Gx$  := CalculaMedias( $Gx(1), \dots, Gx(N)$ );
(07)  $Gy$  := CalculaMedias( $Gy(1), \dots, Gy(N)$ );
(08)  $M_{F_M}$  := CalculaMódulos( $Gx, Gy$ );
(09)  $M_{F_A}$  := CalculaÁngulos( $Gx, Gy$ );
(10)  $M_F$  := [ $M_{F_M}, M_{F_A}$ ];
```

---

$i = 1, \dots, N$ , donde la posición real de fovea ( $C_{F_i}$ ) es conocida para todas las imágenes. Inicialmente, se aplica una etapa de preprocesado siguiendo los pasos ya descritos en la subsección 3.2.1 y en el siguiente orden: primero, se realiza una normalización del canal verde,  $I_G$ , de la imagen de entrada. Solo se usará este canal porque existe evidencia en la literatura de que es el que ofrece mayor contraste para segmentar la fovea [Xu, 2010, Gegundez et al., 2013, Raja and Ravichandran, 2014]. Seguidamente, el canal verde normalizado es transformado a la resolución de trabajo ( $K_0$ ) y, si procede, reflejado horizontalmente (junto con su *ground truth* de fovea) para unificar el tipo de ojo. A continuación, un **bucle-for** recorre todas las imágenes y, para cada una de ellas, se aplica el operador gradiente a la región de píxeles abarcada por una subventana,  $W_F$ , centrada en fovea y de tamaño  $L \times L$ , obteniendo dos matrices,  $Gx(i)$  y  $Gy(i)$ ,  $i = 1, \dots, N$ , con el valor de las componentes  $x$  e  $y$  de los vectores gradiente. Una vez procesadas todas las imágenes, se calcula el valor promedio de cada componente,  $Gx$  y  $Gy$ . Finalmente, las matrices  $M_{F_M}$  y  $M_{F_A}$  almacenan el resultado de calcular el módulo y el ángulo, respectivamente, de los vectores gradiente promedio a partir del valor de sus dos componentes promedio.

Para construir el modelo de fovea, se utilizó un subconjunto de 546 imágenes MESSIDOR, sin ninguna patología macular. El valor elegido para el tamaño de la subventana fue  $L = 27 px$  (a resolución  $K_0$ ). Estos valores son obtenidos considerando que el radio típico de fovea está entre  $\frac{2}{3}$  y  $\frac{2}{4}$  de un  $R_{DO}$  [Schwiegerling, 2004, Gegundez et al., 2013, Xu, 2010], donde el valor de  $R_{DO}$ , a resolución  $K_0$ , se obtiene de la Ec. (3.2).

La Fig. 3.11 muestra información de interés relacionada con el modelo de fovea obtenido. En particular, se muestra la matriz de ángulos (Fig. 3.11.c) y módulos (3.11.d) del campo vectorial promedio resultante (3.11.b) de aplicar el operador gradiente a la región de fovea abarcada por  $\mathbf{W}_F$  (Fig. 3.11.a) de las diferentes imágenes utilizadas en la construcción del modelo.

### Cómo evaluar una instancia del modelo de fovea

Obtenido el modelo de fovea, la pregunta es cómo evaluar una instancia del modelo, es decir, cómo evaluar una subventana de imagen de dimensiones  $L \times L$  y centrada en un punto  $\hat{\mathbf{C}}_F$ , dada una imagen de retina,  $\mathbf{I}$ , es decir:

$$\widehat{\mathbf{M}}_F = \hat{f}_F(\mathbf{I}, \hat{\mathbf{C}}_F) \quad (3.16)$$

donde  $\hat{f}_F$  es la función de evaluación. El Algoritmo 3.4 muestra los pasos necesarios para la implementación de  $\hat{f}_F$ . Así, en primer lugar, en la imagen de entrada, se realiza el mismo preprocesado que el que se realizó para crear el modelo de fovea, conservando los parámetros  $K_0$ ,  $ojo$  y  $L$  el mismo valor que el de los respectivos parámetros usados para crearlo. El resto del algoritmo corre paralelo al del bucle externo del Algoritmo 3.3, con la única diferencia de que los vectores gradientes que se evalúan corresponden al área abarcada por la subventana centrada en  $\hat{\mathbf{C}}_F$ . Finalmente, para evaluar cuánto se aproxima la instancia al modelo, bastará calcular:

$$error_F = f_{errF}(\mathbf{M}_F, \widehat{\mathbf{M}}_F) \quad (3.17)$$

donde  $f_{errF}$  puede ser cualquier función de error, elegida por el usuario, que permita calcular la discrepancia entre  $\mathbf{M}_F$  y  $\widehat{\mathbf{M}}_F$ .

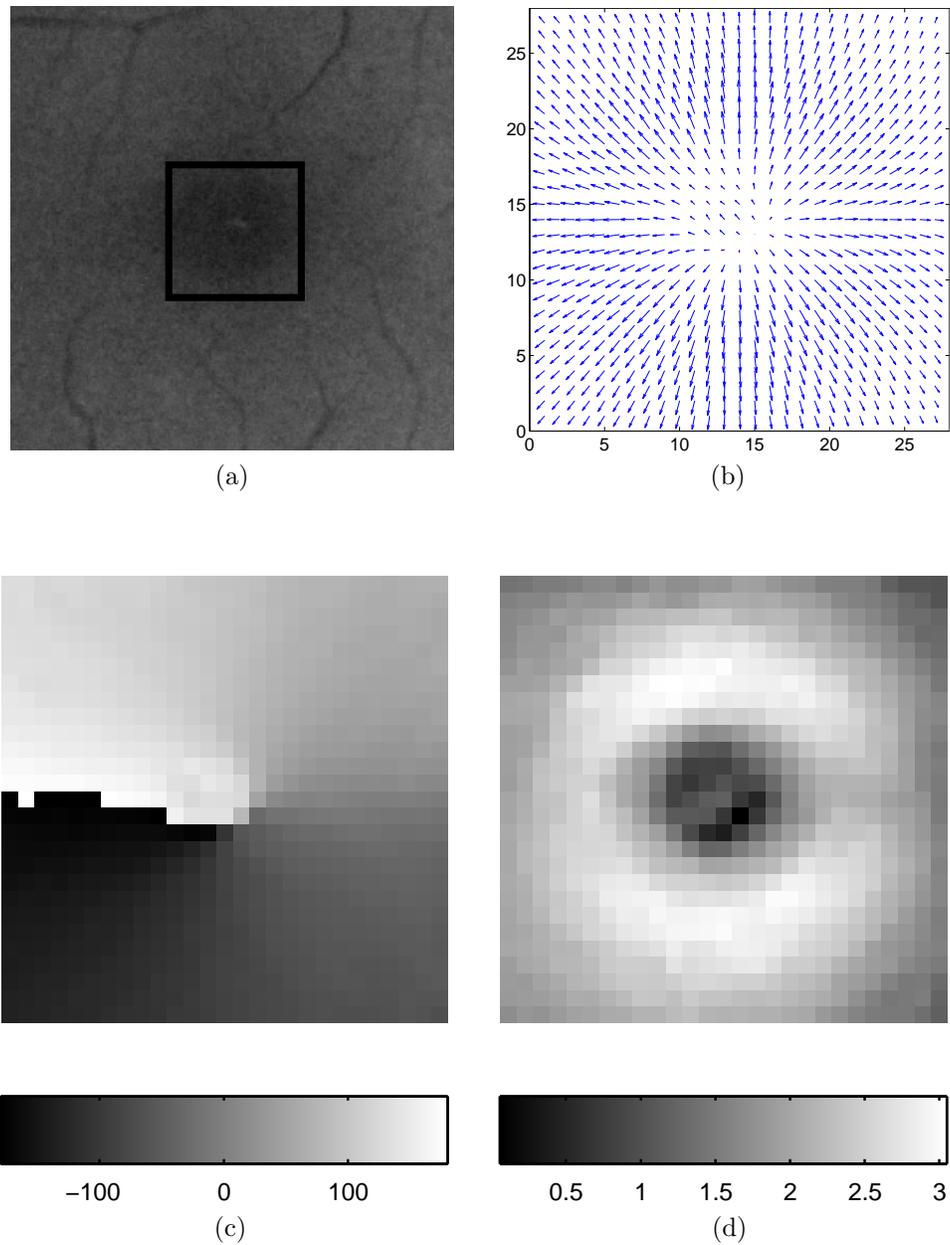


Figura 3.11: Modelo de fovea ( $\mathbf{M}_F$ ) basado en el cálculo de ángulos y módulos del campo vectorial promedio resultante de aplicar el operador gradiente a la región de fovea de diferentes imágenes MESSIDOR: (a) Vista parcial de una imagen de retina (canal verde) en la que se muestra la subventana  $\mathbf{W}_F$  centrada en fovea, (b) campo vectorial gradiente promedio, (c) ángulos del campo vectorial gradiente promedio, (d) módulos del campo vectorial gradiente promedio.

---

**Algoritmo 3.4** Pseudocódigo para evaluar una instancia del modelo de fóvea.

---

**Entradas**

$\mathbf{I}$ , imagen de entrada

$K_0$ , *ojo*,  $L$ , mismos valores que los definidos en Algoritmo 3.3

$\widehat{\mathbf{C}}_F$ , centro de subventana,  $\widehat{\mathbf{W}}_F$ , a evaluar como instancia de fóvea

**Salidas**

$\widehat{\mathbf{M}}_F$ , evaluación de la instancia del modelo de fóvea

---

- (01)  $[\mathbf{I}_G] := \text{Preprocesa}(\mathbf{I}, \text{ojo}, K_0)$ ;  
(02)  $\widehat{\mathbf{W}}_F = \text{ConstruyeSubventana}(\mathbf{I}_G, \widehat{\mathbf{C}}_F, L)$ ;  
(03)  $[\widehat{\mathbf{G}}_x, \widehat{\mathbf{G}}_y] := \text{CalculaGradiente}(\widehat{\mathbf{W}}_F)$ ;  
(04)  $\widehat{\mathbf{M}}_{FM} := \text{CalculaMódulos}(\widehat{\mathbf{G}}_x, \widehat{\mathbf{G}}_y)$ ;  
(05)  $\widehat{\mathbf{M}}_{FA} := \text{CalculaÁngulos}(\widehat{\mathbf{G}}_x, \widehat{\mathbf{G}}_y)$ ;  
(06)  $\widehat{\mathbf{M}}_F = [\widehat{\mathbf{M}}_{FM}, \widehat{\mathbf{M}}_{FA}]$ ;
- 

### 3.2.4. Modelo relacional DO-Fóvea

El objetivo del modelo relacional DO-fóvea es construir un nuevo modelo que aglutine los modelos de DO y de fóvea construidos anteriormente y que incluya el conocimiento relacional disponible entre ambas estructuras anatómicas. Así, es conocido que la distancia media entre el centro geométrico del DO y el centro de la fóvea,  $\overline{D}(\mathbf{C}_{DO}, \mathbf{C}_F)$ , es 5 veces el radio del DO [Sinthanayothin et al., 1999], es decir:

$$\overline{D}(\mathbf{C}_{DO}, \mathbf{C}_F) = 5R_{DO} \quad (3.18)$$

Nótese que el valor de  $\overline{D}(\mathbf{C}_{DO}, \mathbf{C}_F)$  es una constante que puede calcularse a la resolución de trabajo,  $K_0$ , a partir de las Ec. (3.18) y (3.2). Por tanto, la distancia entre el centro del DO y de la fóvea de cualquier imagen puede expresarse como:

$$D(\mathbf{C}_{DO}, \mathbf{C}_F) = \overline{D}(\mathbf{C}_{DO}, \mathbf{C}_F) + \varepsilon \quad (3.19)$$

donde  $\varepsilon$  es un incremento de distancia, positivo o negativo, que representa la posible desviación de la distancia media entre DO y fóvea. De otro lado, también es conocido que, normalmente, los centros de DO y de fóvea no están alineados horizontalmente, sino que existe un cierto ángulo,  $\varphi$ , positivo o negativo, entre el segmento que une ambos centros y

la horizontal [Williams and Wilkinson, 1992, Rohrschneider, 2004, Jonas et al., 2015, Kao et al., 2014].

Combinando toda esta información relacional con los modelos de DO y de fóvea obtenidos, se podría construir un modelo geométrico relacional DO-fóvea tal y como muestra la Fig. 3.12, construido a partir de los modelos de DO y de fóvea:

$$\mathbf{M}_{DO-F} = f_{DO-F}(\mathbf{M}_{DO}, \mathbf{M}_F) \quad (3.20)$$

Sin embargo, no es necesario construir explícitamente este modelo para poder evaluar una instancia del mismo  $(\widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho}, \widehat{\varphi}, \widehat{\varepsilon})$ , es decir, la evaluación de una elipse de DO situada en la imagen,  $(\widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho})$ , cuyo centro está a una distancia del centro de fóvea  $\overline{D}(\mathbf{C}_{DO}, \mathbf{C}_F) + \varepsilon$  y formando el segmento que une ambos centros un ángulo con la horizontal de  $\widehat{\varphi}$ . Efectivamente, a partir de la terna  $(\widehat{\mathbf{C}}_{DO}, \widehat{\varphi}, \widehat{\varepsilon})$ , resulta inmediato calcular el valor de las componentes de  $\widehat{\mathbf{C}}_F$ :

$$\widehat{C}_{Fx} = \widehat{C}_{DOx} + (\overline{D}(C_{DO}, C_F) + \widehat{\varepsilon}) \cdot \cos(\widehat{\varphi}) \quad (3.21)$$

$$\widehat{C}_{Fy} = \widehat{C}_{DOy} + (\overline{D}(C_{DO}, C_F) + \widehat{\varepsilon}) \cdot \sin(\widehat{\varphi})$$

Por tanto, la evaluación de una instancia del modelo relacional puede hacerse como:

$$\begin{aligned} \widehat{\mathbf{M}}_{DO-F} = \widehat{f}_{DO-F}(\mathbf{I}, \widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho}, \widehat{\varphi}, \widehat{\varepsilon}) = \\ [ \widehat{f}_{DO}(\mathbf{I}, \widehat{\mathbf{C}}_{DO}, \widehat{a}, \widehat{b}, \widehat{\rho}, \widehat{\mathbf{C}}_F), \widehat{f}_F(\mathbf{I}, \widehat{\mathbf{C}}_F) ] \end{aligned} \quad (3.22)$$

donde  $\widehat{f}_{DO}$  y  $\widehat{f}_F$  se calculan según Ec. (3.13) y (3.16), respectivamente.

Finalmente, para evaluar cuánto se aproxima la instancia al modelo, bastará calcular:

$$\begin{aligned} error_{DO-F} = f_{errDO-F}(\mathbf{M}_{DO-F}, \widehat{\mathbf{M}}_{DO-F}) = \\ \omega_{DO} \cdot f_{errDO}(\mathbf{M}_{DO}, \widehat{\mathbf{M}}_{DO}) + \omega_F \cdot f_{errF}(\mathbf{M}_F, \widehat{\mathbf{M}}_F) \end{aligned} \quad (3.23)$$

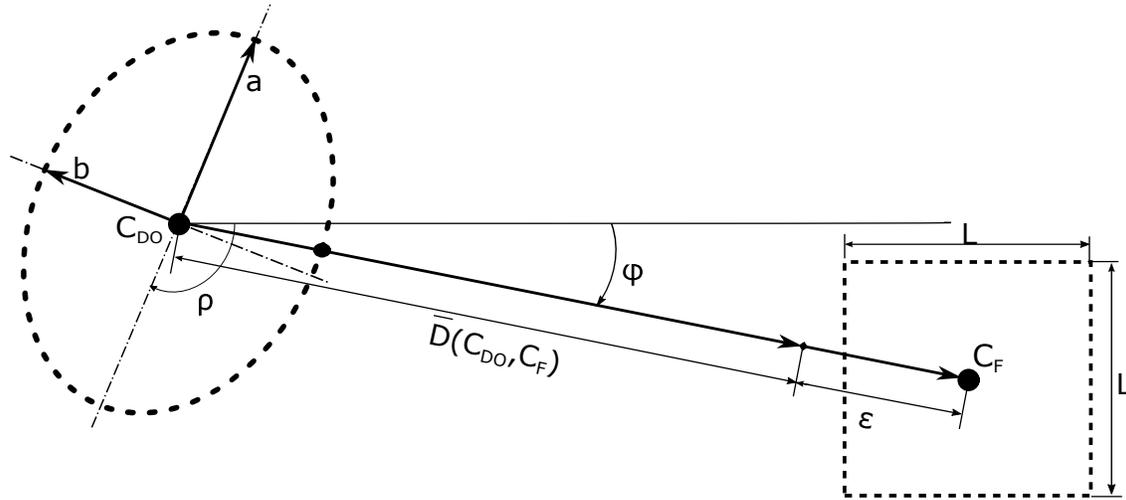


Figura 3.12: Representación esquemática del modelo relacional DO-fóvea ( $M_{DO-F}$ ).

donde  $\omega_{DO}$  y  $\omega_F$  son dos parámetros de ponderación elegidos por el usuario y  $f_{errDO}$  y  $f_{errF}$  se calculan según las Ec. (3.15) y (3.17), respectivamente.

### 3.2.5. Uso de Algoritmos Evolutivos

La potencia del modelo relacional planteado en la sección anterior radica en permitir plantear el problema de la búsqueda y segmentación de las dos estructuras retinianas, DO y fóvea, como un problema de optimización consistente en minimizar la expresión (3.23), teniendo en cuenta la transformación (3.21), es decir:

$$\min_{(C_{DO}, a, b, \rho, \varphi, \epsilon)} error_{DO-F} \quad (3.24)$$

Dado que el problema así definido depende de la optimización de siete parámetros reales ( $C_{DO}$  se desdobra en dos componentes), es imposible abordarlo mediante fuerza bruta. Por tanto, para resolverlo, se propone el uso de un algoritmo evolutivo. En esta sección, se describen los diferentes componentes necesarios para abordar un problema de optimización mediante este tipo de algoritmos: representación de los individuos, definición de la función *fitness* e instanciación del algoritmo evolutivo.

### 3.2. SEGMENTACIÓN DE ESTRUCTURAS RETINIANAS

Tabla 3.2: Descripción de cada uno de los genes codificados en un cromosoma. Todos los valores mostrados en los dominios de definición están calculados a la resolución de trabajo ( $K_0 = 25 \text{ px/mm}$ ). Adicionalmente, las medias y desviaciones estándar,  $\mu_a \pm \sigma_a = 24.3 \pm 2.2$  y  $\mu_b \pm \sigma_b = 22.2 \pm 2.2$ , fueron calculadas a partir de las 1200 imágenes MESSIDOR.

#Gen	Gen	Descripción	Dominio de definición
1	$C_{DOx}$	Coordenada-x de $\mathbf{C}_{DO}$	$[\text{mín}(x), \text{máx}(x)] / \exists y p(x, y) \in \text{retina}$
2	$C_{DOy}$	Coordenada-y de $\mathbf{C}_{DO}$	$[\text{mín}(y), \text{máx}(y)] / \exists x p(x, y) \in \text{retina}$
3	$a$	Semieje mayor de la elipse	$[\mu_a - 3\sigma_a, \mu_a + 3\sigma_a] = [17.7, 30.9]$
4	$b$	Semieje menor de la elipse	$[\mu_b - 3\sigma_b, \mu_b + 3\sigma_b] = [15.6, 28.8]$
5	$\rho$	Ángulo( $a, eje_X$ )	$[0, 180]^\circ$
6	$\varphi$	Ángulo( $\overline{\mathbf{C}_{DO} \mathbf{C}_F}, eje_X$ )	$[-25, 25]^\circ$
7	$\varepsilon$	Offset respecto a $\overline{D}(\mathbf{C}_{DO}, \mathbf{C}_F)$	$[-R_{DO}, R_{DO}] = [-23.1, 23.1]$

#### Representación de individuos

Cada individuo (cromosoma) codificará una instancia del modelo relacional DO-fóvea definido en la subsección 3.2.4. Así, un cromosoma se define como una cadena lineal de longitud fija, donde cada variable (gen) se asocia a cada uno de los parámetros del modelo relacional. La Tabla 3.2 muestra la correspondencia entre cada gen del cromosoma y los parámetros del modelo, así como el dominio de definición de cada uno de ellos. En particular, para la elección del dominio de definición del parámetro  $\varphi$  se usó un criterio similar al usado en [Kao et al., 2014].

#### Función Fitness

La función *fitness* nos permitirá evaluar cada individuo. Su construcción consistirá en definir dos funciones de error,  $f_{errDO}$  (ver Ec. (3.15)) y  $f_{errF}$  (ver Ec. (3.17)), que nos permitan

evaluar el error final dado por la Ec. (3.23). Aquí proponemos las dos funciones siguientes:

$$f_{errDO} = \frac{\sum_{i=1}^{N_p} (\mathbf{máx}_{DO}(i) - \widehat{\mathbf{M}}_{DO}(i))}{\sum_{i=1}^{N_p} (\mathbf{máx}_{DO}(i))} \quad (3.25)$$

$$f_{errF} = \frac{\sum_{i=1}^L \sum_{j=1}^L \mathbf{M}_{FM}(i, j) \cdot |(\mathbf{M}_{FA}(i, j) - \widehat{\mathbf{M}}_{FA}(i, j))|}{180 \cdot \sum_{i=1}^L \sum_{j=1}^L \mathbf{M}_{FM}(i, j)} \quad (3.26)$$

donde  $\mathbf{máx}_{DO}(i) = \max\{\mathbf{M}_{DO}(i), 1 - \mathbf{M}_{DO}(i)\}$ .

Obsérvese que el valor devuelto por ambas funciones está normalizado en el intervalo  $[0, 1]$ . Así, para DO, el peor caso ( $f_{errDO} = 1$ ) corresponde a una elipse en la que ninguno de sus puntos es punto frontera o vaso, es decir,  $\forall i \widehat{\mathbf{M}}_{DO}(i) = 0$  y, el mejor caso ( $f_{errDO} = 0$ ), cuando todos los puntos de la elipse son clasificados correctamente, es decir,  $\forall i \widehat{\mathbf{M}}_{DO}(i) = \mathbf{máx}_{DO}(i)$ . De otro lado, para fóvea, el peor caso ( $f_{errF} = 1$ ) se da cuando el error absoluto evaluado en cada sumando del numerador devuelve siempre el error máximo de ángulo, es decir,  $180^\circ$  y, el mejor caso ( $f_{errF} = 0$ ), cuando la matriz de ángulos del modelo y la instancia coinciden, es decir,  $\forall ij \widehat{\mathbf{M}}_{FA}(i, j) = \mathbf{M}_{FA}(i, j)$ .

### Conocimiento del dominio y restricciones

El conocimiento relacional intra-estructura (CRintra-e) se puede utilizar también para establecer restricciones sobre el valor de los parámetros de un modelo o sobre otros parámetros derivados. En nuestro caso, analizamos el conocimiento estadístico de uno de los parámetros directos del modelo ( $\rho$ ) y de dos parámetros derivados: área de la elipse ( $A$ ) y diferencia entre semiejes de la elipse ( $\lambda$ ). La elección de estos tres parámetros está basada en el siguiente conocimiento del dominio: (a) la orientación del eje mayor de las elipses *gold estándar* tiende a ser vertical ( $\rho \approx 90^\circ$ ); (b) Dado que el contorno de la excavación papilar puede presentar un patrón parecido al del DO, se utiliza la restricción del área para descartar posibles falsos positivos asociados a esta estructura; (c) el semieje mayor de la elipse tiene que ser siempre mayor que el semieje menor y, además, la excentricidad de las elipses *gold estándar* tiende a ser pequeña ( $\lambda > 0$ ). Analizando las 1200 imágenes MESSIDOR a la escala de trabajo ( $K_0$ ),

### 3.2. SEGMENTACIÓN DE ESTRUCTURAS RETINIANAS

Tabla 3.3: Cálculo del intervalo de restricción  $(X_o, X_f)$  para que el 95 % de los valores extraídos de la distribución asociada a cada variable estén dentro de dicho intervalo.

Variable ( $X$ )	Función de densidad de probabilidad	$(X_o, X_f)$
$A = \pi ab$	Normal ( $\mu = 1712.3px^2$ , $\sigma = 317.1px^2$ )	(2346.5, 1078.1)
$\rho = \text{Ángulo}(a, eje_X)$	Normal ( $\mu = 90.6^\circ$ , $\sigma = 22.8^\circ$ )	(45.0, 136.2)
$\lambda = (a - b)$	Gamma ( $A = 2.6817$ , $B = 0.7979$ )	(0.39, 5.36)

la Tabla 3.3 muestra las funciones de densidad de probabilidad de cada una de estas variables (parámetros) y los extremos del intervalo  $(X_o, X_f)$  para que el 95 % de los valores extraídos de cada una de las distribuciones obtenidas estén dentro de dicho intervalo.

La idea es utilizar cada uno de estos intervalos para definir restricciones que permitan penalizar elipses genéticas que se alejen del valor de dichos intervalos. La forma más inmediata de añadir dichas restricciones al problema de optimización es incorporarlas dentro de la función que evalúa el error del modelo de DO, es decir, en  $f_{errDO}$ . Para ello, se definen tres funciones de penalización,  $P(\rho)$ ,  $P(A)$  y  $P(\lambda)$ , siguiendo el esquema dado por la función genérica (3.27), donde los extremos del intervalo  $(X_o, X_f)$  vienen detallados en la Tabla 3.3 para cada valor posible de  $X \in \{\rho, A, \lambda\}$ .

$$P(X) = \begin{cases} 0 & \text{si } X \in (X_o, X_f) \\ 10 \cdot \frac{\min\{|X-X_o|, |X-X_f|\}}{(X_f-X_o)} & \text{en otro caso} \end{cases} \quad (3.27)$$

Dado que estamos abordando un problema de minimización, el problema de minimizar  $f_{errDO}$ , teniendo en cuenta las restricciones dadas por los respectivos intervalos  $(X_o, X_f)$  para  $\rho$ ,  $A$  y  $\lambda$ , se transformará en el problema de minimizar una nueva función dada por:

$$f_{errDO}^{(r)} = f_{errDO} + P(\rho) + P(A) + P(\lambda) \quad (3.28)$$

Finalmente, la función *fitness* a minimizar por el algoritmo evolutivo, incluyendo las

restricciones, resulta ser:

$$F = \omega_{DO} \cdot f_{errDO}^{(r)} + \omega_F \cdot f_{errF} \quad (3.29)$$

### Algoritmo Evolutivo

Aquí se propone el uso de *differential evolution* (DE) [Storn and Price, 1997] como algoritmo evolutivo (AE). Varias razones motivan esta decisión. En primer lugar, DE es simple y directo de implementar, proporcionando resultados prometedores. En segundo lugar, DE maneja de forma natural individuos expresados como vectores reales de dimensión fija. Tercero, la complejidad del espacio de parámetros de control en DE es bajo comparado con algunos de los más competitivos optimizadores de parámetros reales [Das and Suganthan, 2011]. Por último, DE ya ha sido utilizado con éxito para abordar la tarea de segmentación en imágenes para diferentes dominios médicos [Aslantas and Tunckanat, 2007, Mesejo et al., 2013, Saraswat et al., 2013, Ugolotti et al., 2013, Mesejo et al., 2016]. No obstante, teniendo en cuenta el esquema usado para la codificación de cada individuo (vector real de dimensión fija), otros AEs podrían haberse usado como, por ejemplo, algoritmos genéticos [Holland, 1975, Michalewicz, 1996], estrategias evolutivas [Beyer and Schwefel, 2002] o CMA-ES (del inglés, *Covariance Matrix Adaptation Evolution Strategy*) [Hansen, 2006].

Básicamente, DE sigue las etapas típicas de un AE (inicialización, mutación, recombinación y selección), pero se caracteriza por una implementación particular de la mutación. Dentro de las diferentes variantes de DE [Price et al., 2005], aquí hemos seleccionado la variante DE estándar (ver Algoritmo 3.5). Así, para cada individuo de la población,  $\mathbf{x}_i$ , la mutación consiste en la construcción de un nuevo individuo dado por:

$$\mathbf{y} = \mathbf{x}_p + F \cdot (\mathbf{x}_q - \mathbf{x}_r) \quad (3.30)$$

donde  $\mathbf{x}_p$ ,  $\mathbf{x}_q$  y  $\mathbf{x}_r$  son tres individuos seleccionados de la población de forma aleatoria, sin reemplazamiento y diferentes a  $\mathbf{x}_i$ . La recombinación se realiza entre las diferentes componentes de  $\mathbf{x}_i$  e  $\mathbf{y}$  con una probabilidad  $CR$ . Finalmente, si el *fitness* del hijo resultante,  $\mathbf{y}$ , es mejor que el de  $\mathbf{x}_i$ , entonces el primero reemplaza al segundo en la población. La notación

---

**Algoritmo 3.5** Pseudocódigo del algoritmo “DE/rand/1/bin”.

---

**Inputs:**

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ , función *fitness* a minimizar  
 $CR \in [0, 1]$ , probabilidad de cruce  
 $F \in [0, 2]$ , peso diferencial  
 $NP$ , tamaño de población

**Output:**

$\mathbf{x}^*$ , el mejor individuo de la población final ( $\mathbf{P}$ )

---

```

(01) Inicializa aleatoriamente  $\mathbf{P} = \{\mathbf{x}_1, \dots, \mathbf{x}_{NP}\}$ 
(02) until (termination criterion)
(03) for each  $\mathbf{x}_i \in \mathbf{P}$ 
(04)   Elegir al azar  $\{p, q, r\} \in \{1, \dots, NP\} / i \neq p, i \neq q, i \neq r, p \neq q, p \neq r, q \neq r$ 
(05)   Elegir al azar  $R \in \{1, \dots, n\}$ 
(06)   for each  $j \in \{1, \dots, n\}$ 
(07)      $r_j \equiv U(0, 1)$ 
(08)     if  $r_j < CR$  or  $j = R$  then  $y_j = x_{pj} + F \cdot (x_{qj} - x_{rj})$  else  $y_j = x_{ij}$  end-if
(09)   end-for
(10)   if  $f(\mathbf{y}) < f(\mathbf{x}_i)$  then  $\mathbf{x}_i = \mathbf{y}$  end-if %Se asume un problema de minimización
(11) end-for
(12) end-until

```

---

“DE/x/y/z” se interpreta de la siguiente forma: la etiqueta “/x/” hace referencia a cómo se selecciona el individuo que interviene en la mutación haciendo el papel de  $\mathbf{x}_p$  (en el algoritmo DE estándar, aleatoriamente “ $x = rand$ ”); la etiqueta “/y/” denota cuantas parejas de individuos intervienen en el cálculo de la diferencia mostrada en la Ec. 3.30 (en el algoritmo DE estándar, “ $y = 1$ ”); y la etiqueta “/z/” indica el tipo de recombinación usada (en el algoritmo DE estándar, binomial, es decir, “ $z = bin$ ”, tal y como se muestra en las líneas 6-9 del Algoritmo 3.5).

### 3.3. Segmentación con información de detección

Como ya se comentó la sección 3.1, el resultado de detección proporciona una posición aproximada del centro de la estructura anatómica que se desea buscar. Por tanto, con el objetivo de mejorar el coste computacional y/o los resultados de segmentación, la salida del método de detección podría ser utilizada como una entrada adicional al método de segmentación.

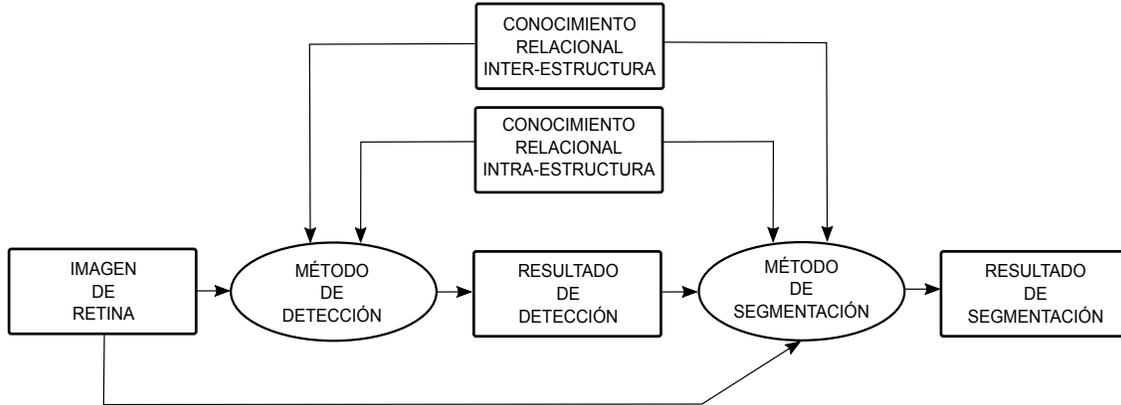


Figura 3.13: Acoplamiento del método de detección al método de segmentación

La Fig. 3.13 muestra un diagrama de bloques del acoplamiento de ambos métodos. El método de segmentación es casi idéntico al descrito en la sección 3.2, existiendo una única diferencia: el algoritmo evolutivo crea una población inicial utilizando como semilla los resultados de detección, a diferencia del método original en el que la población inicial se creaba aleatoriamente y sin ningún conocimiento previo.

El algoritmo 3.6 muestra el proceso de generación de la población inicial ( $\mathbf{P}_0$ ). Básicamente, la idea consiste en utilizar como semilla los centros de las estructuras obtenidos en la etapa de detección para generar aleatoriamente una población intermedia,  $\mathbf{P}$ , de  $N$  individuos situados en la vecindad de dichos centros y, finalmente, seleccionar una subpoblación  $\mathbf{P}_0$  con los  $N_0$  mejores individuos contenidos en  $\mathbf{P}$ , siendo  $N$  y  $N_0$ , con  $N \gg N_0$ , parámetros proporcionados por el usuario.

A continuación, se hace una descripción más detallada del proceso de generación de  $\mathbf{P}_0$ . Así, a partir de la solución del método de detección, formada por los centros de las tres estructuras,  $(\mathbf{C}_{DO}, \mathbf{C}_{HV}, \mathbf{C}_F)$ , se generan dos parejas,  $(\mathbf{C}_{DO}, \mathbf{C}_F)$  y  $(\mathbf{C}_{HV}, \mathbf{C}_F)$ . Dado que los centros  $\mathbf{C}_{DO}$  y  $\mathbf{C}_{HV}$  siempre están muy próximos (ver sección 3.1.3), estos pueden considerarse casi equivalentes y, por tanto, las dos parejas mencionadas pueden verse como dos soluciones aproximadas de la pareja real de centros de DO y fovea. No obstante, antes de usar ambas parejas como semillas DO-fovea, es preciso chequear si la solución de detección se generó a partir de algún blob de DO, fovea o HV no contenido en alguna tripleta o pareja válidas (ver paso 7 del pseudocódigo descrito en la sección 3.1.3), en cuyo caso no queda garantizado

### 3.3. SEGMENTACIÓN CON INFORMACIÓN DE DETECCIÓN

---

parte del CRinter-e empleado en detección (ver condiciones (b-c) y (f-g) del paso 5 del pseudocódigo descrito en la sección 3.1.3). Es por esto que se utiliza la función booleana `ChequeaValidezPareja` para chequear si cada una de las dos parejas ( $\mathbf{C}_{DO}$ ,  $\mathbf{C}_F$ ) y ( $\mathbf{C}_{HV}$ ,  $\mathbf{C}_F$ ) verifica o no dicho conocimiento relacional. A continuación, dependiendo del resultado de este chequeo, se realizan cuatro acciones diferentes. Si ambas parejas son válidas, el 50% de la población  $\mathbf{P}$  se generará tomando como semilla una de ellas y, el 50% restante, con la otra pareja. Si solo una de ellas es válida, el 100% de la población  $\mathbf{P}$  se generará tomando como semilla la pareja válida. En última instancia, si ninguna es válida, el 100% de la población  $\mathbf{P}$  se generará de forma aleatoria y  $\mathbf{P}_0$  se obtendrá seleccionando los  $N_0$  mejores individuos de  $\mathbf{P}$ .

La forma de generar los individuos a partir de las dos parejas de semillas (suponiendo que al menos una de ellas es válida), se muestra en las líneas 19-22 del algoritmo 3.6. Así, uno de los individuos de la población  $\mathbf{P}$  se construye codificando directamente la semilla ( $\mathbf{SC}_{DO_1}$ ,  $\mathbf{SC}_F$ ) en los genes respectivos de su cromosoma (ver Tabla 3.2). Concretamente, la función `GeneraIndividuo` (línea 19) almacena la semilla  $\mathbf{SC}_{DO_1}$  directamente en los genes  $C_{DO_x}$  y  $C_{DO_y}$ , siendo codificada la semilla  $\mathbf{SC}_F$  de forma indirecta en los genes  $\varphi$  y  $\varepsilon$ , mediante la aplicación de la transformación indicada en la Ec.(3.21). El resto de genes,  $a$ ,  $b$  y  $\rho$ , se selecciona de forma aleatoria dentro de sus respectivos dominio de definición (ver Tabla 3.2). Este proceso se repite para la generación de un segundo individuo (línea 20), pero tomando como semilla la pareja ( $\mathbf{SC}_{DO_2}$ ,  $\mathbf{SC}_F$ ).

El resto de individuos se construye a partir de la función `GeneraIndividuoGauss` (líneas 21-22), cuya misión es generar centros de parejas DO-fóvea a partir de una distribución gaussiana 2D con centro en la pareja usada como semilla ( $\mathbf{SC}_{DO_i}$ ,  $\mathbf{SC}_F$ ), con  $i = 1, 2$ , y con una desviación típica igual a  $\frac{1}{3}R_{DO}$  (con el objeto de que el 99% de la población generada esté dentro de un círculo con centro en la semilla y radio igual a  $1R_{DO}$ ). Una vez generados los centros DO-fóvea mediante la distribución gaussiana, el proceso de generación de los respectivos genes del cromosoma de cada individuo es análogo al empleado en la función `GeneraIndividuo`, ya comentada en el párrafo anterior.

Una vez creada toda la población  $\mathbf{P}$ , se realiza un proceso de chequeo (líneas 24-34) para

comprobar si el valor de los genes  $C_{DO_x}$ ,  $C_{DO_y}$ ,  $\varepsilon$  y  $\varphi$  de cada individuo cumplen las restricciones asociadas al CRinter-e usado en la tarea de segmentación (la validez de los genes  $a$ ,  $b$  y  $\rho$  queda garantizada por construcción). Para ello, se utiliza la función `ChequeaCondiciones`, una función booleana que devuelve *true* si, dado el individuo  $i$ -ésimo, alguna de las siguientes condiciones es cierta:

1. El centro generado de DO,  $(C_{DO_x}, C_{DO_y})$ , se encuentra fuera de la zona de retina o si la distancia de este a la semilla de DO,  $\mathbf{SC}_{DO}$ , es mayor que  $p$  veces un  $R_{DO}$ . El valor de  $p$  se toma igual a  $\frac{3}{2}$ , considerando así un umbral de distancia algo menor a la distancia máxima admisible ( $2R_{DO}$ ).
2. El centro generado de fovea,  $(C_{F_x}, C_{F_y})$ , se encuentra fuera de la zona de retina o si la distancia de este a la semilla de fovea,  $\mathbf{SC}_F$ , es mayor que  $p$  veces un  $R_{DO}$ . El valor de  $p$  usado es idéntico al caso anterior.
3. El valor de  $\varepsilon$ , obtenido a partir de  $(C_{F_x}, C_{F_y})$  y la Ec.(3.21), está fuera de su dominio de definición  $[-R_{DO}, R_{DO}]$  (ver Tabla 3.2).
4. El valor de  $\varphi$ , obtenido a partir de  $(C_{F_x}, C_{F_y})$  y la Ec.(3.21), está fuera de su dominio de definición  $[-25, 25]^\circ$  (ver Tabla 3.2).

Por tanto, el individuo que incumpla alguna de estas cuatro restricciones se generará nuevamente y este proceso se repetirá hasta que la función `ChequeaCondiciones` devuelva un valor *false*. Una vez que todos los individuos son válidos, se obtiene la población  $\mathbf{P}$ . Finalmente, la población inicial  $\mathbf{P}_0$  se obtendrá seleccionando los  $N_0$  mejores individuos de  $\mathbf{P}$  (línea 35). Seleccionada  $\mathbf{P}_0$ , el algoritmo evolutivo se aplica de la misma forma que se describe en la sección 3.2.

### 3.3. SEGMENTACIÓN CON INFORMACIÓN DE DETECCIÓN

---

**Algoritmo 3.6** Pseudocódigo para generar una población inicial a partir de la solución obtenida por el método de detección.

---

**Entradas**

$[\mathbf{C}_{DO}, \mathbf{C}_{HV}, \mathbf{C}_F]$ , tripleta solución del método de detección

$N$ , Tamaño población temporal

$N_0$ , Tamaño población inicial ( $N \gg N_0$ )

**Salida**

$\mathbf{P}_0 := \{[\mathbf{C}_{DO_1}, a_1, b_1, \rho_1, \varphi_1, \varepsilon_1], \dots, [\mathbf{C}_{DO_{N_0}}, a_{N_0}, b_{N_0}, \rho_{N_0}, \varphi_{N_0}, \varepsilon_{N_0}]\}$ , Población inicial

---

```
(01)  $e_1 := \text{ChequeaValidezPareja}(\mathbf{C}_{DO}, \mathbf{C}_F);$ 
(02)  $e_2 := \text{ChequeaValidezPareja}(\mathbf{C}_{HV}, \mathbf{C}_F);$ 
(03) switch ( $e_1, e_2$ )
(04)   case (1, 1):
(05)      $\mathbf{SC}_{DO_1} := \mathbf{C}_{DO};$ 
(06)      $\mathbf{SC}_{DO_2} := \mathbf{C}_{HV};$ 
(07)   case (1, 0):
(08)      $\mathbf{SC}_{DO_1} := \mathbf{C}_{DO};$ 
(09)      $\mathbf{SC}_{DO_2} := \mathbf{C}_{DO};$ 
(10)   case (0, 1):
(11)      $\mathbf{SC}_{DO_1} := \mathbf{C}_{HV};$ 
(12)      $\mathbf{SC}_{DO_2} := \mathbf{C}_{HV};$ 
(13)   default:
(14)      $\mathbf{P} := \text{GeneraPoblaciónAleatoriamente}(N);$ 
(15)      $\mathbf{P}_0 := \text{Selecciona\_}N_0\text{\_mejores}(\mathbf{P}, N_0);$ 
(16)     return( $\mathbf{P}_0$ )
(17) end-switch
(18)  $\mathbf{SC}_F := \mathbf{C}_F$ 
(19)  $\mathbf{P}[1] := \text{GeneraIndividuo}(\mathbf{SC}_{DO_1}, \mathbf{SC}_F);$ 
(20)  $\mathbf{P}[N] := \text{GeneraIndividuo}(\mathbf{SC}_{DO_2}, \mathbf{SC}_F);$ 
(21)  $\text{SubP}_1 := \text{GeneraIndividuoGauss}(\mathbf{SC}_{DO_1}, \mathbf{SC}_F, (N-2)/2);$ 
(22)  $\text{SubP}_2 := \text{GeneraIndividuoGauss}(\mathbf{SC}_{DO_2}, \mathbf{SC}_F, (N-2)/2);$ 
(23)  $\mathbf{P} := \text{SubP}_1 \cup \text{SubP}_2;$ 
(24) for  $i := 2$  to  $N - 1$ 
(25)    $\text{cond} := \text{ChequeaCondiciones}(\mathbf{P}[i]);$ 
(26)   while ( $\text{cond}$ )
(27)     if ( $i \leq \frac{N}{2}$ )
(28)        $\mathbf{P}[i] := \text{GeneraIndividuoGauss}(\mathbf{SC}_{DO_1}, \mathbf{SC}_F, 1);$ 
(29)     else
(30)        $\mathbf{P}[i] := \text{GeneraIndividuoGauss}(\mathbf{SC}_{DO_2}, \mathbf{SC}_F, 1);$ 
(31)     end-if
(32)      $\text{cond} := \text{ChequeaCondiciones}(\mathbf{P}[i]);$ 
(33)   end-while
(34) end-for
(35)  $\mathbf{P}_0 := \text{Selecciona\_}N_0\text{\_mejores}(\mathbf{P}, N_0);$ 
(36) return( $\mathbf{P}_0$ )
```

---

# Capítulo 4

## Resultados experimentales y discusión

En este capítulo se describen los experimentos de detección y segmentación propuestos, mostrando los resultados obtenidos y realizando una discusión de los mismos. Por una parte, en la sección 4.1 se detallan los experimentos relacionados con detección. Luego, en la sección 4.2, se muestran los resultados obtenidos con el método de detección implementado, tanto en la base de imágenes MESSIDOR como en otras bases de imágenes públicas. A continuación, la sección 4.3 muestra una detallada discusión de resultados sobre estos experimentos, comparándolos con otros métodos competitivos existentes en la literatura. Adicionalmente, se analizan otros aspectos de interés relacionados con las características del método implementado. Por otra parte, para la segmentación, se sigue una estructura similar a la descrita para detección. En primer lugar, en la sección 4.4, se describen los experimentos de segmentación realizados. Posteriormente, en la sección 4.5, se presentan los resultados obtenidos con el método de segmentación implementado en las bases de imágenes MESSIDOR y ONHSD. Seguidamente, en la sección 4.6, se discuten los resultados obtenidos en detalle, comparándolos con otros métodos de segmentación competitivos, así como otras propiedades relacionadas con el método implementado. Finalmente, en la sección 4.7 se detallan aquellos experimentos dirigidos a evaluar el método de segmentación simultánea con información de detección, incluyendo los resultados obtenidos y la discusión de los mismos. Es importante señalar que, en la discusión de resultados de cada uno de los tres métodos implementados, también se anali-

zan los resultados de aquellos experimentos dirigidos a evidenciar las hipótesis planteadas en esta tesis.

### 4.1. Diseño de experimentos de detección

En esta sección se describen los diferentes experimentos llevados a cabo para evaluar el método de detección. También se detalla los valores asociados a la configuración de parámetros elegida para el método, así como los criterios usados para evaluar los resultados de detección obtenidos.

Los experimentos diseñados para evaluar diferentes aspectos del método de detección son los siguientes. El **primer experimento** consiste en evaluar el método en la base de imágenes MESSIDOR. Son varias las razones que hacen de este conjunto de imágenes el más utilizado actualmente en la evaluación tanto de métodos de detección como de segmentación: es de acceso público, posee una gran cantidad de imágenes (1200) con una amplia casuística (imágenes normales o con diferentes grados de patologías, presencia o no de artefactos ruidosos, diferentes condiciones de iluminación, etc.) y, finalmente, y no menos importante, el *ground truth* de DO y mácula de sus imágenes también es accesible públicamente. El **segundo experimento** va dirigido a evaluar la influencia de diferentes signos de patología en las prestaciones del método. Para esto, se hace uso también de MESSIDOR, teniendo en cuenta que es conocido el grado de retinopatía diabética y el riesgo de edema macular de cada una de sus imágenes. El **tercer experimento** es similar al primero, pero utilizando otras bases de imágenes públicas (DIARETDB1, ONHSD y DRIONS). Así, a diferencia de MESSIDOR, cada una de estas bases de imágenes contiene una menor cantidad de imágenes. No obstante, dos fueron los objetivos principales a la hora de utilizarlas: (i) analizar la robustez del método en otras bases de imágenes diferentes a MESSIDOR, conservando la misma configuración de parámetros; y (ii) poder comparar los resultados obtenidos de detección de HV con otros métodos, dado que, hasta donde se sabe, esta estructura no ha sido evaluada en MESSIDOR. El **cuarto experimento** va dirigido a analizar la influencia del carácter simultáneo de la detección y de la cantidad de CRinter-e inyectado en los resultados de detección. Por

tanto, el objetivo de este experimento es refutar o añadir evidencia a las dos primeras hipótesis de investigación planteadas en esta tesis. El **quinto experimento** persigue analizar las prestaciones del método automático de detección del tipo de ojo (izquierdo o derecho) en imágenes de retina que fue propuesto en la sección 2.3.1. Por último, el **sexto experimento** va dirigido a intentar sacar provecho de los resultados obtenidos por el método de detección, proponiendo dos nuevas estrategias de refinamiento dirigidas a mejorar sus prestaciones.

En cuanto al valor de los parámetros de los que depende el método de detección, se elige una configuración que se mantiene fija para todos los experimentos descritos anteriormente. Así, el valor final elegido para la resolución estándar fue  $K_0 = 9 \text{ px/mm}$ . Conocido  $K_0$ , el valor del factor de escala,  $\alpha$ , necesario para transformar cada imagen a la resolución  $K_0$ , se calcula utilizando la Ec. (3.1). El valor  $K_0$  fue elegido lo suficientemente pequeño como para eliminar artefactos ruidosos de la imagen original y lo suficientemente grande como para no obstaculizar la detección de las estructuras buscadas. En cualquier caso, este valor no es crítico, pudiendo haber elegido otro valor parecido. Para todos los experimentos, el número de blobs candidatos fue  $N_{DO} = 9$  para DO ( $n_{do} = 3$ , es decir, tres blobs por cada canal RGB),  $N_M = 9$  para los blobs de mácula ( $n_m = 3$ , es decir, tres por cada canal RGB), y  $N_{HV} = 4$  para los blobs de HV ( $n_{hv} = 2$ , es decir, dos por cada tipo de ojo). Estos valores se obtuvieron experimentalmente, teniendo en cuenta que valores más bajos podrían no incluir el blob solución en el conjunto de blobs candidatos e, inversamente, valores más altos podrían aumentar el tiempo de procesamiento del método.

Los criterios utilizados para evaluar los resultados de detección de las diferentes estructuras coinciden con los usados normalmente en la literatura. Así, al igual que se hace en la mayoría de los trabajos [Tobin et al., 2007, Yu et al., 2011, Gegundez et al., 2013, Giachetti et al., 2014], consideramos un criterio exitoso para la detección de DO si la distancia entre los centros de DO estimado y real es menor o igual a un radio de DO ( $1R_{DO}$ ). En cuanto a la mácula, la mayor parte de los trabajos de investigación han utilizado el centro de la misma como punto de referencia. Así, la detección de mácula será exitosa si la distancia entre los centros de mácula estimado y real es menor o igual que  $p$  veces  $1R_{DO}$ . En este caso, dependiendo del enfoque considerado en la literatura, el valor de  $p$  puede ser diferente. Por ejemplo,

en [Yu et al., 2011] se considera  $p = 1/2$ ,  $p = 1$  en [Giachetti et al., 2013, Aquino, 2014], o  $p = 2$  en [Tobin et al., 2007]. Finalmente, en relación con el HV, se considera una detección exitosa si la distancia entre el centro del HV estimado y el centro real de DO es menor o igual que  $1R_{DO}$ . Elegimos el centro de DO real como referencia porque este está siempre muy cerca del centro del HV (ver Fig. 1.1) y, dado que en la tarea de detección buscamos solo una aproximación a la posición de la estructura real, esta pequeña discrepancia no será significativa a la hora de evaluar los resultados de detección del HV.

## 4.2. Resultados de detección

Los resultados de detección se presentan de forma paralela a los experimentos definidos en la sección anterior.

### 4.2.1. Evaluación en la base de imágenes MESSIDOR

La Tabla 4.1 muestra una comparativa del método con otros trabajos en términos de detección solo de DO y/o mácula en la base de imágenes MESSIDOR. Como puede comprobarse en la tabla, también se muestran los resultados de detección de trabajos que utilizan bases de imágenes privadas. Hay que tener en cuenta que, inicialmente, no existían bases de imágenes públicas y, por tanto, cada grupo de investigación evaluaba su método con su propia base de imágenes. En estos casos, la comparativa debería tener en cuenta esta circunstancia. En cualquier caso, creemos que es interesante mostrar los resultados de algunos de estos métodos de detección pioneros por la relevancia de los resultados obtenidos.

Por otra parte, también es importante evaluar cómo puede afectar la existencia de diferentes patologías oculares al rendimiento del método propuesto. Así, las Tablas 4.2 y 4.3 muestran una comparativa de nuestros resultados de detección de DO y mácula con otros enfoques en la literatura, considerando diferentes grados de retinopatía y diferentes riesgos de edema macular en la base de imágenes MESSIDOR, respectivamente.

Adicionalmente, resulta interesante mencionar que si analizamos la posibilidad de fallo del método en dos estructuras simultáneamente, solo se obtuvo una imagen en la que se produjo

Tabla 4.1: Comparativa de los resultados de detección de DO y mácula (expresada en %) obtenidos en MESSIDOR. Los resultados con los que se compara se muestran tal y como se presentan en los trabajos donde se publicaron. Para hacer más justa la comparación, los valores mostrados para [Niemeijer et al., 2009, Niemeijer et al., 2007] corresponden al resultado de considerar los conjuntos de imágenes normales y patológicas como un único conjunto. Las imágenes pueden ser normales (norm) o patológicas (patol).

Método	Base de imágenes	#Imágenes (norm+patol)	DO	HV	Mácula ( $p$ )		
					$(\frac{1}{2}R_{DO})$	$(1R_{DO})$	$(2R_{DO})$
[Fleming et al., 2007]	Privada	1056	98.40	-	-	96.50	98.50
[Tobin et al., 2007]	Privada	345 (18+327)	90.40	-	-	-	92.50
[Niemeijer et al., 2007]	Privada	600 (500+100)	97.67	-	-	94.00	-
[Niemeijer et al., 2009]	Privada	549 (449+100)	98.23	-	-	95.38	-
[Aquino et al., 2010]	MESSIDOR	1200 (546+654)	99.00	-	-	-	-
[Yu et al., 2011]	MESSIDOR	1200 (546+654)	98.33	-	95.00	-	-
[Gegundez et al., 2013]	MESSIDOR	1200 (546+654)	-	-	96.08	96.92	97.83
[Giachetti et al., 2013]	MESSIDOR	1136 (533+603)	99.66	-	-	<b>99.10</b>	-
[Aquino, 2014]	MESSIDOR	1136 (533+603)	-	-	91.28	98.24	99.56
[Kao et al., 2014]	MESSIDOR	1200 (546+654)	-	-	<b>97.83</b>	98.17	-
[Marin et al., 2015]	MESSIDOR	1200 (546+654)	<b>99.75</b>	-	-	-	-
[Girard et al., 2016]	MESSIDOR	1200 (546+654)	-	-	94.00	98.00	-
[Dashtbozorg et al., 2016]	MESSIDOR	1200 (546+654)	<b>99.75</b>	-	93.75	98.87	99.58
[Al-Bander et al., 2018]	MESSIDOR	1200 (546+654)	97.00	-	91.14	96.60	-
<b>Este trabajo</b>	MESSIDOR	1200 (546+654)	99.33	<b>97.92</b>	96.25	98.83	<b>99.75</b>

## 4.2. RESULTADOS DE DETECCIÓN

Tabla 4.2: Comparativa de los resultados de detección de DO (expresado en %) obtenidos en MESSIDOR para diferentes grados de retinopatías y riesgos de edema macular. Los resultados con los que se compara se muestran tal y como se presentan en los trabajos publicados.

Método	Grado de retinopatía					Riesgo de edema macular			
	0	1	2	3	Cualq.	0	1	2	Cualq.
[Yu et al., 2011]	98.71	99.34	99.19	96.16	98.33	98.39	<b>100</b>	97.26	98.33
[Giachetti et al., 2013]	<b>99.62</b>	<b>100</b>	<b>100</b>	<b>99.23</b>	<b>99.66</b>	<b>99.69</b>	<b>100</b>	<b>99.35</b>	<b>99.66</b>
<b>Este trabajo</b>	99.08	99.35	<b>100</b>	99.21	99.33	99.28	<b>100</b>	99.34	99.33

Tabla 4.3: Comparativa de los resultados de detección de mácula (expresada en %) obtenidos en MESSIDOR para diferentes grados de retinopatías y riesgos de edema macular. Los resultados con los que se compara se muestran tal y como se presentan en los trabajos publicados, excepto los mostrados para [Gegundez et al., 2013], que fueron obtenidos de [Giachetti et al., 2013]. El método propuesto se evalúa para dos valores de  $p$  diferentes ( $1R_{DO}$  y  $\frac{1}{2}R_{DO}$ ), a fin de poder compararse con diferentes métodos.

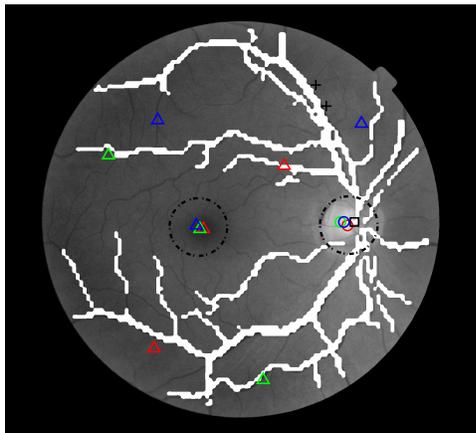
Método	$p$	Grado de retinopatía					Riesgo de edema macular			
		0	1	2	3	Cualq.	0	1	2	Cualq.
[Giachetti et al., 2013]	$1R_{DO}$	<b>99.80</b>	<b>100</b>	<b>99.20</b>	<b>98.70</b>	<b>99.10</b>	<b>99.40</b>	<b>100</b>	<b>96.90</b>	<b>99.10</b>
[Gegundez et al., 2013]	$1R_{DO}$	99.20	98.70	99.10	95.40	98.20	98.80	98.60	95.80	98.20
<b>Este trabajo</b>	$1R_{DO}$	99.63	<b>100</b>	98.79	96.46	98.83	99.18	<b>100</b>	96.03	98.83
[Yu et al., 2011]	$\frac{1}{2}R_{DO}$	<b>97.78</b>	99.35	95.95	85.77	95.00	96.19	98.67	85.71	95.00
<b>Este trabajo</b>	$\frac{1}{2}R_{DO}$	97.62	<b>100</b>	<b>97.57</b>	<b>89.76</b>	<b>96.25</b>	<b>97.33</b>	<b>100</b>	<b>87.42</b>	<b>96.25</b>

este tipo de fallo, obteniendo una detección errónea de DO y mácula. De otro lado, no se obtuvo ningún caso en el que las tres estructuras obtenidas como solución correspondieran a tres falsos positivos.

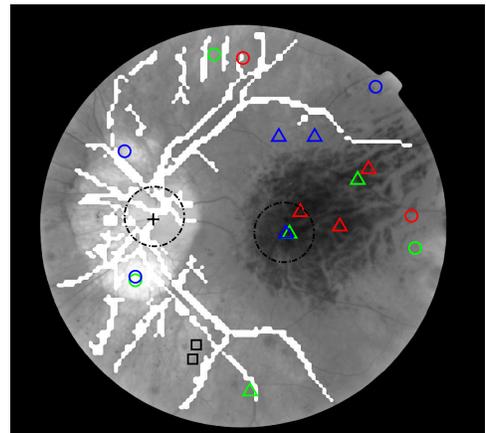
Finalmente, las Figs. 4.1 y 4.2 muestran cuatro ejemplos con salidas intermedias del proceso de detección en imágenes MESSIDOR. Específicamente, las subfiguras 4.1(a, c, e) muestran la consecución de una detección exitosa de las tres estructuras y las subfiguras 4.1(b, d, f) muestran una secuencia en la que solo el DO se detecta erróneamente. Por otra parte, las subfiguras 4.2(a, c, e) muestran una secuencia en la que solo la mácula es errónea y, finalmente, las subfiguras 4.2(b, d, f) muestran una secuencia en la que solo el HV es erróneo. Concretamente y, para ambas figuras, la primera fila muestra el conjunto de blobs candidatos para cada estructura (ver sección 3.1.2). La segunda fila muestra el resultado del filtrado de blobs de acuerdo con los pasos 1-3 del algoritmo de selección de tripletas (ver sección 3.1.3). La tercera fila muestra los blobs solución (paso 6 o 7 del algoritmo de selección de tripletas). Los círculos, triángulos, cuadrados y cruces representan el DO, la mácula, el HV derecho e izquierdo, respectivamente (cada color está asociado con su respectivo canal RGB). Las circunferencias discontinuas están centradas en el *ground truth* de DO y de mácula.

#### 4.2.2. Evaluación en otras bases de imágenes públicas

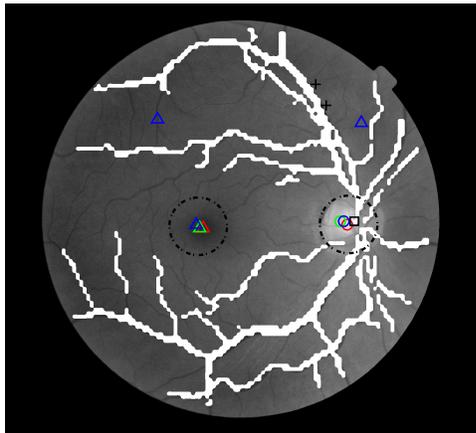
Con el fin de proporcionar evidencia sobre la robustez de la metodología presentada, se utilizaron otras bases de imágenes públicas: DIARETDB1, ONHSD y DRIONS. En este caso, es importante señalar que, con estas bases de imágenes, el método fue evaluado utilizando la misma configuración de parámetros que se utilizó con la base de imágenes MESSIDOR, tal y como se describió en la sección 4.1. La Tabla 4.4 muestra los resultados obtenidos. Hay que tener en cuenta que la base de imágenes DRIONS tiene la peculiaridad de no mostrar el área macular en la mayoría de sus imágenes, por lo que nuestro método solo se evaluó para DO y HV. A diferencia de lo que ocurre en MESSIDOR, donde no se pudo encontrar ningún trabajo que evaluara la detección del HV en dicha base de imágenes, esto sí fue posible hacerlo en DIARETDB1.



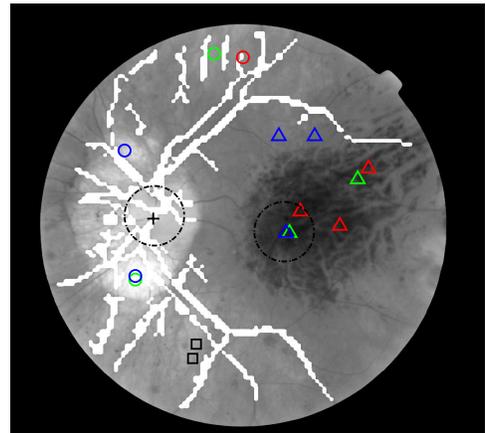
(a)



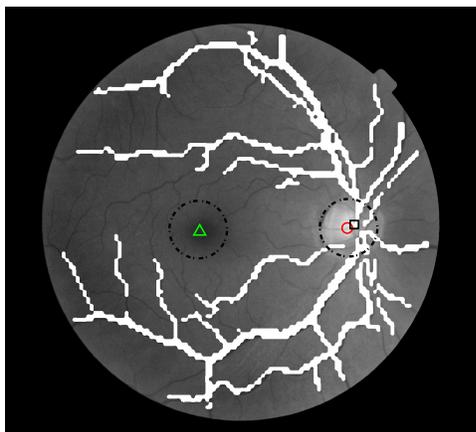
(b)



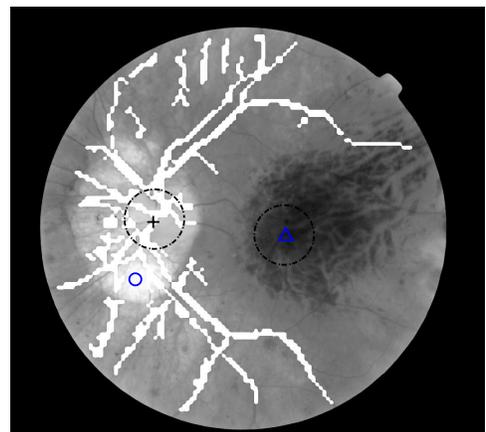
(c)



(d)



(e)



(f)

Figura 4.1: Dos ejemplos que muestran los resultados del proceso de detección en diferentes imágenes MESSIDOR: éxito en las tres estructuras detectadas (1<sup>a</sup> columna); éxito en mácula/HV y fallo en DO (2<sup>a</sup> columna). Las filas muestran distintas etapas del proceso de detección (la última muestra al resultado final).

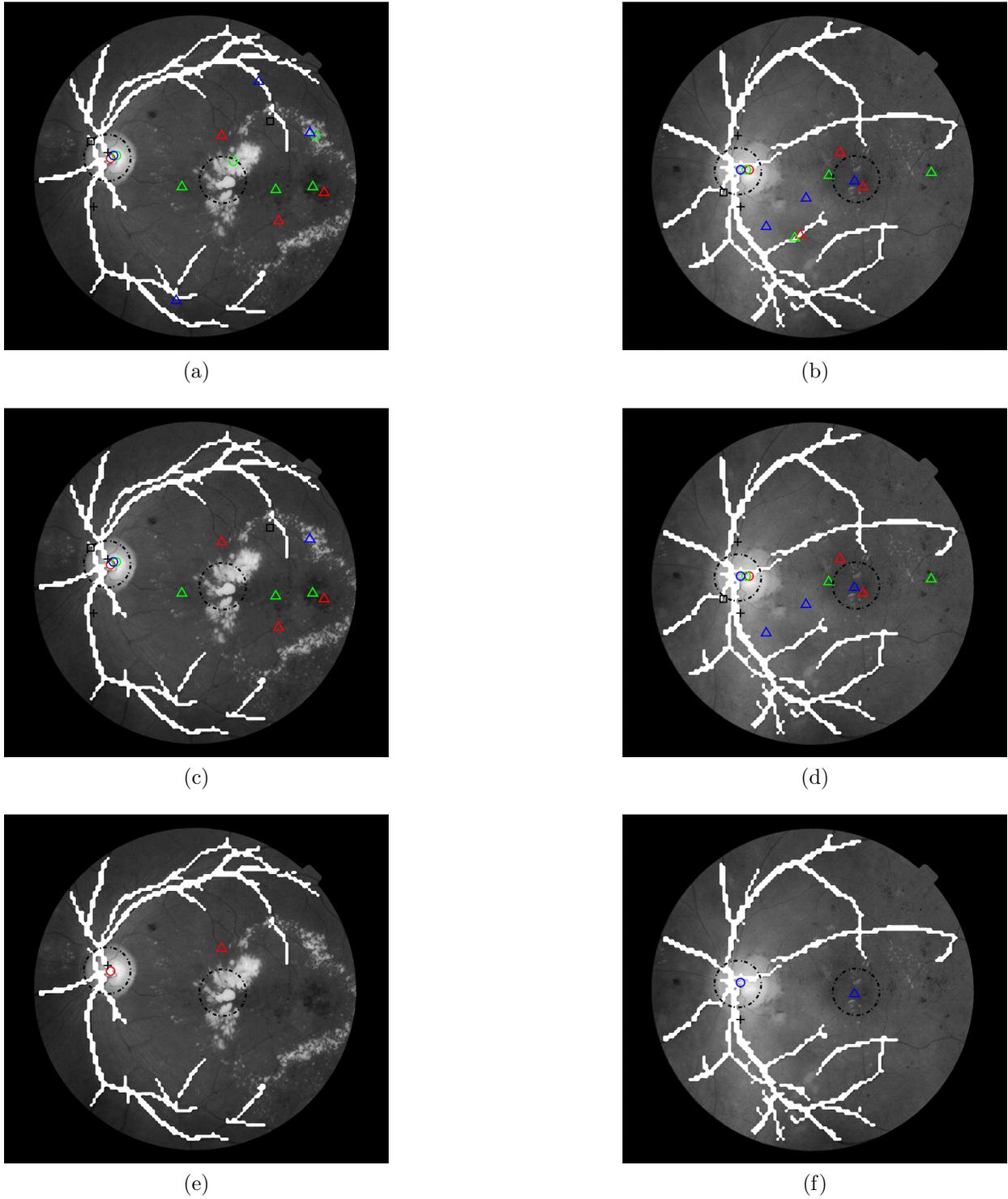


Figura 4.2: Dos ejemplos que muestran los resultados del proceso de detección en diferentes imágenes MESSIDOR: éxito en DO/HV y fallo en macula (1<sup>a</sup> columna); éxito en DO/mácula y fallo en HV (2<sup>a</sup> columna). Las filas muestran distintas etapas del proceso de detección (la última muestra el resultado final).

## 4.2. RESULTADOS DE DETECCIÓN

---

Tabla 4.4: Comparativa de los resultados de detección de DO, mácula y HV (expresados en%) obtenidos por el método propuesto en otras bases de imágenes públicas diferentes a MESSIDOR.

Método	Base de datos	#Imágenes	DO	Mácula ( $p=1R_{DO}$ )	HV
[Mahfouz and Fahmy, 2010]	DIARETDB1	89	97.75	-	<b>97.75</b>
[Welfer et al., 2011]	DIARETDB1	89	-	92.13	-
[Qureshi et al., 2012]	DIARETDB1	89	<b>97.79</b>	98.74	-
[Cao et al., 2013]	DIARETDB1	89	97.75	-	<b>97.75</b>
[Aquino, 2014]	DIARETDB1	89	-	94.38	-
[Kao et al., 2014]	DIARETDB1	89	-	92.13	-
[Mittal and Sivaswamy, 2015]	DIARETDB1	89	97.75	96.00	-
[Medhi and Dandapat, 2016]	DIARETDB1	89	-	95.51	-
[Xiong and Li, 2016]	DIARETDB1	89	97.75	-	-
<b>Este trabajo</b>	DIARETDB1	89	96.63	<b>100</b>	<b>97.75</b>
[Lowell et al., 2004]	ONHSD	87	<b>100</b>	-	-
<b>Este trabajo</b>	ONHSD	87	<b>100</b>	<b>98.85</b>	<b>97.70</b>
[Molina and Carmona, 2011]	DRIONS	110	<b>100</b>	-	-
<b>Este trabajo</b>	DRIONS	110	<b>100</b>	-	<b>100</b>

### 4.3. Discusión de resultados de detección

En toda la bibliografía manejada, no se han encontrado trabajos que aborden la detección y evaluación simultánea de las tres estructuras detectadas aquí (DO, mácula y HV). Por lo tanto, solo se pueden comparar los resultados de la metodología de detección con otros enfoques donde dos o menos estructuras anatómicas son detectadas y evaluadas, ya sea de forma aislada, secuencial o simultánea. A continuación se realiza una discusión de los resultados obtenidos.

#### 4.3.1. Base de imágenes MESSIDOR

Los índices de éxito obtenidos con el método propuesto son competitivos en relación con otros métodos aplicados a la base de imágenes MESSIDOR (ver Tabla 4.1). Concretamente, los resultados de detección de DO y mácula obtenidos por nuestro método son, respectivamente, 99.33% y [96.25% ( $p = 1/2$ ), 98.83% ( $p = 1$ ), 99.75% ( $p = 2$ )], resaltando que el porcentaje obtenido por nuestro método para detección de mácula, con  $p = 2$ , supera a todos los métodos de la comparativa.

Si se usan todos los resultados disponibles, independientemente del número de estructuras detectadas por cada aproximación y sin tener en cuenta el caso de mácula con el que obtenemos las mejores prestaciones ( $p = 2$ ), nuestros resultados son solo ligeramente superados por dos métodos [Marin et al., 2015, Giachetti et al., 2013]. Así, en [Giachetti et al., 2013], el DO y la mácula se detectan, respectivamente, con una tasa de éxito de 99.66% y 99.10% ( $p = 1$ ), pero los resultados de detección de la mácula solo están disponibles para  $p = 1$  y, además, se obtuvieron para 1136 de las 1200 imágenes de MESSIDOR (se desconoce el porqué no se utilizaron las 64 restantes en la evaluación). Por otra parte, en [Marin et al., 2015], la tasa de éxito de DO es 99.75%, pero es la única estructura detectada y evaluada.

Aunque la detección secuencial de DO y mácula se aborda en [Kao et al., 2014], solo se proporcionan resultados para la mácula [97.83% ( $p = 1/2$ ), 98.17% ( $p = 1$ )], obteniendo la tasa de éxito más alta para  $p = 1/2$ , comparada con todos los métodos utilizados en nuestra comparativa (nuestra tasa de éxito para  $p = 1/2$  es la segunda mejor).

### 4.3. DISCUSIÓN DE RESULTADOS DE DETECCIÓN

---

De otro lado, en relación a los trabajos que, como en nuestro caso, realizan detección simultánea, se aplican los siguientes comentarios. En primer lugar, todos ellos abordan la detección simultánea de solo dos estructuras: mácula y DO. Así, en [Dashtbozorg et al., 2016], se obtienen buenos resultados de detección de DO y mácula (99.75 % y [93.75 % ( $p = 1/2$ ), 98.87 % ( $p = 1$ ), 99.58 % ( $p = 2$ )]), pero la comparativa arroja distintos resultados dependiendo de lo que se compare. Así, las tasas de éxito para DO y mácula ( $p = 1$ ) son ligeramente superiores a las obtenidas por nuestro método, pero ligeramente inferiores para mácula con  $p = 1/2$  y  $p = 2$  (96.25 % y 99.75 % respectivamente). En [Girard et al., 2016], se detectan parejas candidatas de DO-mácula, para luego seleccionar la mejor pareja. Sin embargo, sus resultados de detección de mácula [94.0 % ( $p = 1/2$ ), 98.0 % ( $p = 1$ )] son ligeramente más bajos que los nuestros y sus resultados de DO, al ser solo evaluados desde el punto de vista de la tarea de segmentación, no pueden ser comparados. Por último, en [Al-Bander et al., 2018], se obtienen resultados relativamente discretos tanto en DO como en mácula, concretamente 97.00 % y [91.14 % ( $p = 1/2$ ), 96.60 % ( $p = 1$ )], respectivamente, siendo superados por nuestro trabajo y por casi todos los trabajos en esta comparativa. Como ventaja de este último trabajo podemos citar la velocidad de ejecución del proceso de detección.

En relación al resto de trabajos usados en la comparativa, nuestros resultados de detección de DO y mácula igualan o superan ligeramente a los obtenidos en [Yu et al., 2011, Gegundez et al., 2013] y superan a los obtenidos en [Tobin et al., 2007, Fleming et al., 2007, Niemeijer et al., 2009, Niemeijer et al., 2007].

De otro lado, la influencia del grado de retinopatía diabética (GRD) y del riesgo de edema macular (REM) en los resultados de detección de DO y mácula de nuestro método, no resulta ser significativa, como se puede observar en las Tablas 4.2 y 4.3, respectivamente. Estos resultados proporcionan evidencia de la robustez del método en presencia de distractores asociados a diferentes patologías. Así, el GRD se expresa en función del número de microaneurismas ( $\mu A$ ), del número de hemorragias ( $H$ ) y la presencia de neovascularización ( $NV = 1$ ):

- “0 (Normal)”  $\rightarrow (\mu A = 0) \wedge (H = 0)$
- “1”  $\rightarrow (0 < \mu A \leq 5) \wedge (H = 0)$
- “2”  $\rightarrow [(5 < \mu A < 15) \vee (0 < H < 5)] \wedge (NV = 0)$
- “3”  $\rightarrow (\mu A \geq 15) \vee (H \geq 5) \vee (NV = 1)$

De otro lado, el REM se expresa en términos de la presencia o ausencia de exudados duros ( $hE$ ):

- “0 (Sin riesgo)”  $\rightarrow$  “No visible  $hE$ ”
- “1”  $\rightarrow \min\_distancia(mácula, hE) > 2R_{DO}$
- “2”  $\rightarrow \min\_distancia(mácula, hE) \leq 2R_{DO}$

En relación al DO, los éxitos de detección son siempre superiores al 99 %, independientemente del GRD o del REM. En cualquier caso, los errores en DO suelen aparecer en imágenes con atrofias peripapilares avanzadas o estafilomas posteriores (signos no considerados para evaluar el GRD o el REM), tal y como, por ejemplo, ocurre en la Fig. 4.1(f). Con respecto a la mácula ( $p = 1$ ), el método obtiene un porcentaje de éxito superior o igual al 98.8 % en imágenes con  $GRD = \{0,1,2\}$  o  $REM = \{0,1\}$ . Sin embargo, este porcentaje cae al 96.46 % en aquellas imágenes con un elevado número de microaneurismas o hemorragias ( $GRD = 3$ ) y desciende al 96.03 % en aquellas imágenes con grandes exudados duros en el área macular ( $REM = 2$ ) como, por ejemplo, ocurre en la Fig. 4.2(e). Por tanto, la existencia de signos de patología macular en su mayor grado de severidad son los que más pueden dificultar el proceso de detección de mácula usando nuestro método.

### 4.3.2. Otras bases de imágenes

En relación a otras bases de imágenes, los resultados de detección obtenidos para las diferentes estructuras son también competitivos y consistentes con los resultados obtenidos en la base de imágenes MESSIDOR (ver Tabla 4.4). En particular, el resultado de detección

### 4.3. DISCUSIÓN DE RESULTADOS DE DETECCIÓN

---

de mácula en DIARETDB1 (100 %,  $p = 1$ ) es el más alto comparado con el resto de métodos. Por otra parte, en cuanto a DRIONS y ONSHD, el porcentaje de acierto en detección de DO es absoluto (100 %), igualando a [Lowell et al., 2004, Molina and Carmona, 2011].

Especial interés reviste el trabajo de [Qureshi et al., 2012], dado que también se aplica una estrategia de detección simultánea, aunque limitado a dos estructuras (DO y mácula). Por tanto, resulta de interés hacer una comparativa para cada estructura y también de forma global, teniendo en cuenta que no todas las bases de imágenes utilizadas por ellos coinciden con las utilizadas por nosotros. De una parte, mejoramos en mácula (100 % vs. 98.74 %) pero somos ligeramente peores en DO (96.63 % vs. 97.79 %) para DIARETDB1. Atendiendo a la tasa de éxito promedio por estructura, y teniendo en cuenta todas las bases de imágenes utilizadas, obtenemos mejores resultados: 99.25 % frente a 98.06 % para DO, y 98.90 % frente a 96.87 % para mácula. El cálculo de los resultados así obtenidos fue realizado a partir de la Ec. 4.1, donde  $BD_i$  es la base de datos  $i$ -ésima,  $tamaño_{BD_i}$  es el número de imágenes en  $BD_i$ ,  $RD_{BD_i}(e)$  es el ratio de detección (expresado en tanto por uno) alcanzado por el método en la  $BD_i$  para la estructura  $e$ , y  $n$  es el número de bases de imágenes diferentes en las cuales  $e$  fue buscada.

$$RD(e) = \frac{\sum_{i=1}^n tamaño_{BD_i} \times RD_{BD_i}(e)}{\sum_{i=1}^n tamaño_{BD_i}} \quad (4.1)$$

Adicionalmente, el trabajo de [Qureshi et al., 2012] es interesante porque incluye un experimento y resultados que también añaden evidencia a la segunda hipótesis de investigación planteada en esta tesis (ver sección 1.2.3): la detección simultánea de varias estructuras permite mejorar los resultados cuando se comparan a los obtenidos con la detección de cada estructura por separado. Así, en el mencionado trabajo, se obtienen porcentajes de detección de DO y mácula de 98.06 % y 96.87 %, respectivamente, cuando la detección es simultánea, y de 96.34 % y 95.53 %, respectivamente, cuando la detección se hace por separado. El hecho de que la tasa de éxito promedio obtenida por nuestro método mejore la obtenida por [Qureshi et al., 2012], podría explicarse por la mayor cantidad de CRinter-e utilizado en nuestro método. Nótese que nuestro método, además del CRinter-e existente entre DO y mácula, también

utiliza el correspondiente a la red de vasos y al HV.

De otro lado, la discusión de los resultados obtenidos para HV es algo más complicada. Por un lado, existen enfoques que combinan la convergencia de los vasos sanguíneos con la detección de otros patrones para finalmente detectar el DO (véase, por ejemplo, [Marin et al., 2015, Niemeijer et al., 2007]). En estos casos, consideramos que son los resultados de detección de DO (no los resultados de detección de HV) los que se deberían utilizar para comparar con los resultados obtenidos en este tipo de enfoques (ver Tabla 4.1). Por otro lado, hay otros trabajos que buscan directamente el centro del HV como aproximación al centro de DO [Mahfouz and Fahmy, 2010, Cao et al., 2013]. En estos casos, los resultados obtenidos sí son comparables directamente con nuestros resultados de detección de HV. Así, la tasa de éxito de HV obtenida con nuestro método de detección (97.75 %) es igual a la obtenida en [Mahfouz and Fahmy, 2010] y [Cao et al., 2013], considerando la base de imágenes DIARETDB1 (ver Tabla 4.4). En cada una de las dos referencias mencionadas, se utilizan tres bases de imágenes más, pero estas son diferentes a las utilizadas en nuestra evaluación. Sin embargo, todavía podemos comparar la tasa promedio de éxito considerando las cuatro bases de imágenes utilizadas en cada estudio y utilizando la Ec. 4.1. Así, el resultado de esta comparativa arrojó que nuestra tasa de éxito promedio (98.0 %) es ligeramente mejor que la obtenida en [Mahfouz and Fahmy, 2010] y [Cao et al., 2013] e igual a 97.0 % y 96.5 %, respectivamente. En cuanto a los casos en los que nuestro método puede presentar alguna dificultad en la detección del HV, se pueden mencionar aquellas imágenes en las que los vasos principales abandonan el centro del DO siguiendo trayectorias rectilíneas hacia arriba y hacia abajo (véase, por ejemplo, la Fig. 4.2(f)), en lugar de seguir el patrón típico de trayectoria curva mostrada en la Fig. 3.6. En estos casos, la marca de detección suele estar unida a un punto de vaso situado fuera del DO, pero normalmente muy cerca de la región superior o inferior del contorno papilar.

A modo de resumen, puede decirse que los buenos resultados obtenidos en DIARETDB1, DRIONS y ONSHD apoyan y refuerzan la hipótesis de que el método propuesto es robusto e independiente de la base de imágenes considerada.

### 4.3.3. Otras consideraciones

Tal y como se mencionó en la sección 4.1, el cuarto experimento propuesto consistía en analizar cómo la detección simultánea y la cantidad de CRinter-e utilizado afecta al proceso de detección. La Tabla 4.5 muestra los resultados obtenidos de este experimento en la base de imágenes MESSIDOR. En el primer estudio, no se utiliza CRinter-e, es decir, cada tipo de estructura se detecta individualmente seleccionando la respuesta más alta del TM, es decir,  $\max(I_{TM}(B_x))$ , donde  $B_x$  es un blob de tipo- $x$ , con  $x \in \{do_i, m_j, hv_k\}$ . Los resultados para este estudio se muestran en la segunda columna y puede verse que son peores que los obtenidos con el método de detección simultánea original (cuarta columna). Estos resultados también añaden evidencia experimental a la segunda hipótesis de investigación planteada en esta tesis (ver sección 1.2.3). En el segundo estudio, se utiliza CRinter-e, pero considerando solamente parejas de estructuras, es decir, no tripletas como en el método original. Aquí, el algoritmo utilizado es el mismo que el mencionado en la sección 3.1.3, pero sin aplicar los pasos 4, 5 y 6. Los resultados se muestran en la tercera columna. Comparados con el experimento anterior, los resultados obtenidos son mejores, es decir, la detección simultánea basada en el uso del conocimiento relacional entre dos estructuras mejora los porcentajes de éxito, pero dichos resultados siguen siendo peores que los del método original donde se utiliza todo el CRinter-e disponible. Los resultados de los dos estudios también añaden evidencia experimental a la primera hipótesis de investigación planteada en esta tesis (ver sección 1.2.3).

La influencia en los resultados de detección de la cantidad de CRintra-e inyectado se puede analizar también considerando como criterio de éxito el que las tres estructuras detectadas correspondan simultáneamente a tres verdaderos positivos. En este caso, los porcentajes de éxito obtenidos en la base de imágenes MESSIDOR fueron los siguientes: 91.91 % si se selecciona la respuesta TM más alta de cada tipo de estructura como solución (no se utiliza CRinter-e), 95.25 % si solo se considera el método de parejas, y 96.16 % si se utiliza el método original (tripletas). Una vez más, se comprueba que el proceso de detección mejora con la detección simultánea y a medida que aumenta la cantidad de CRinter-e inyectado, añadiendo nueva evidencia a nuestras dos primeras hipótesis de investigación.

Tabla 4.5: Tasa de éxito (expresada en %) en la base de imágenes MESSIDOR para cada estructura anatómica, dependiendo de la cantidad de CRinter-e inyectado en el proceso de búsqueda.

Estructura anatómica	Sin CRinter-e $\max\{I_{TM}(\mathbf{B}_x)\}$	CRinter-e por parejas	CRinter-e por tripletas
Disco óptico	98.00	99.16	<b>99.33</b>
Mácula ( $p = 1R_{DO}$ )	97.08	98.75	<b>98.83</b>
Haz vascular	96.58	97.17	<b>97.92</b>
Media	97.22	98.36	<b>98.69</b>

Por otro lado, tal y como se mencionó en la sección 3.1.2, el tipo de *template* que produce la máxima respuesta del TM se puede utilizar como herramienta para conocer de forma automática el tipo de ojo que contiene la imagen analizada (izquierdo o derecho). El quinto experimento planteado en la sección 4.1, orientado a implementar un método para evaluar automáticamente esta tarea, obtuvo un resultado del 100 % de acierto a la hora de identificar el tipo de ojo en las 1200 imágenes MESSIDOR. Por tanto, se comprueba experimentalmente que este procedimiento podría utilizarse aisladamente como un método rápido y simple de identificar el tipo de ojo.

Creemos que una de las ventajas de la metodología de detección propuesta es el uso de dos etapas en las que, en cada una de ellas, se utiliza un tipo de conocimiento diferente (CRintra en la primera e inter-e en la segunda). Esto nos permite usar métodos de detección no excesivamente complejos en la primera etapa porque solo estamos interesados en detectar blobs candidatos para cada estructura, es decir, en esta etapa se admiten falsos positivos entre los blobs candidatos. De otro lado, la segunda etapa solo tiene que usar el CRinter-e para seleccionar la mejor tupla candidata. Por el contrario, hay otros enfoques en la literatura que se plantean como objetivo la detección de una única estructura [Sinthanayothin et al., 1999] o, alternativamente, la detección previa de una estructura que luego se utilizará como referencia en la detección de una segunda [Tobin et al., 2007, Fleming et al., 2007, Yu et al.,

2011, Gegundez et al., 2013]. En ambos casos, los enfoques utilizados son más complejos. En el primer caso, porque la ausencia de CRinter-e tiene que ser compensada con un mayor uso de CRintra-e. En el segundo, porque la detección de la primera estructura también se realiza en ausencia de CRinter-e y, por tanto, comparte las desventajas del primer caso.

Por último, en relación con las condiciones de iluminación, el método propuesto es robusto a cambios de iluminación y solo se ve parcialmente afectado (error en una o dos de las tres estructuras de la solución obtenida) cuando hay problemas de iluminación severos en la imagen.

#### **Coste computacional**

Se pueden analizar también varios aspectos relacionados con la implementación y el coste computacional del método propuesto. Así, el método fue implementado usando lenguaje MATLAB y los experimentos se ejecutaron en un portátil con CPU Intel Core i7-2640M, 2.8GHz y 8GB de RAM. La Tabla 4.6 muestra una comparación con otros métodos en la literatura. Como estos últimos se implementaron en diferentes equipos informáticos, el tiempo de ejecución debería interpretarse en función del número de estructuras detectadas, del hardware y del lenguaje de programación utilizados. En cualquier caso, para realizar una comparación lo más equitativa posible, la tabla mencionada solo muestra los resultados de aquellos trabajos de investigación en los que se detectan dos o más estructuras. Se concluye que nuestro método de detección es rápido, considerando que se detectan cuatro estructuras simultáneamente (DO, mácula, HV y red de vasos principales) y, además, se hace utilizando un lenguaje interpretado. El proceso podría ser aún más rápido si el método se implementara usando un lenguaje compilado.

#### **Refinando la salida del método de detección**

Con el objeto de refinar la salida de detección del método propuesto, en esta sección se describirán dos nuevas estrategias que utilizan como entrada la salida por él producida. Dichas estrategias están asociadas al sexto experimento mencionado en la sección 4.1.

Tabla 4.6: Comparación con métodos competitivos en la literatura considerando el hardware, el lenguaje de programación, el tiempo de cálculo, el número y tipo de estructura detectada: disco óptico (DO), mácula (M), vasos principales (V) y haz vascular (HV).

Método	Hardware	Lenguaje	Tiempo (s)	#Estructuras detectadas (tipo)
[Fleming et al., 2007]	Intel Pentium 4, 2.4GHz	MATLAB & C	120	3 (DO, M, V)
[Niemeijer et al., 2007]	Intel Pentium 4, 3GHz	C++	600	3 (DO, M, V)
[Niemeijer et al., 2009]	Intel Core 2 Duo, 2.83GHz	C++	27.6	3 (DO, M, V)
[Yu et al., 2011]	Intel Xeon W3520, 2.67GHz, 6GB RAM	MATLAB	3.9	2 (DO, M)
[Giachetti et al., 2013]	Intel Core i7 Q 740, 2.93GHz	MATLAB	5	3 (DO, M, V)
[Kao et al., 2014]	Intel Core i7-3610QM, 2.30GHz	C++	[2.6,6.9]	2 (DO, M)
[Marin et al., 2015]	Intel Xeon 3.2GHz, 32GB RAM	MATLAB	$5.42 \pm 0.98$	2 (DO, V)
[Mittal and Sivaswamy, 2015]	2.1GHz, 4GB RAM	MATLAB	[120,160]	2 (DO, M)
<b>Este trabajo</b>	Intel Core i7-2640M, 2.8GHz, 8GB RAM	MATLAB	$4.16 \pm 0.72$	4 (DO, M, V, HV)

### 4.3. DISCUSIÓN DE RESULTADOS DE DETECCIÓN

---

De un lado, cuando el método de detección devuelve una solución basada en tripleta ganadora, nos asegura que dicha solución no tiene inconsistencias o contradicciones en cuanto a las restricciones impuestas basadas en el CRinter-e (ver sección 3.1.3). En otro caso, si el método no devuelve una tripleta, se aplica la búsqueda de parejas o la de blobs individuales para obtener los tres blobs solución. Sin embargo, por construcción del método, los blobs resultantes obtenidos en estos dos últimos casos no cumplirán todas las restricciones. Dicho de otra forma, una solución basada en tripleta reduce la aparición de falsos positivos comparada con una solución obtenida a partir de parejas o blobs individuales. Utilizando esta propiedad, podría utilizarse un módulo adicional a la salida del método de detección que alertara de aquellas imágenes en las que no se produjo una tripleta como solución. Entonces, estas imágenes podrían ser etiquetadas manualmente por un experto o pasadas a otro módulo de detección más preciso y que, probablemente, tendría un coste computacional más alto. Concretamente, del conjunto de imágenes MESSIDOR, solo en 5 de las 1200 imágenes la solución no es obtenida mediante tripleta. Hay que indicar que estas cinco imágenes presentan severos signos de patología y/o importantes problemas de iluminación.

Por otra parte, resulta interesante la propiedad de que el centro del HV está situado muy cerca del centro del DO. En este sentido, la detección del primero puede utilizarse como información redundante para el cálculo del segundo y, por tanto, considerar el valor medio de los centros obtenidos en ambas estructuras como centro del DO. Se propone entonces la siguiente estrategia: si el resultado de detección se produce como resultado de obtener la mejor tripleta, entonces se elige como DO el punto medio de las posiciones resultantes de DO y HV. En otro caso, si se obtiene una o más parejas válidas, se elige la que tenga asociada mayor fitness, de acuerdo con las Ecs. 3.5, 3.6 y 3.7. Se presentan entonces los siguientes casos: (i) Si la pareja ganadora es (DO, HV), entonces se elige el DO como el punto medio entre DO y HV; (ii) Si la pareja ganadora es (HV, mácula), se elige el DO como la posición del HV; y (iii) Si la pareja ganadora es (DO, mácula), el DO conserva su posición. Finalmente, si no existen ni tripletas ni parejas válidas, se elige como DO aquel blob que tenga mayor intensidad TM del conjunto formado por los blobs {DO, HV}. Así, si se aplica esta estrategia al conjunto de imágenes MESSIDOR, se obtiene un 99.75 % de éxito en la detección del DO,

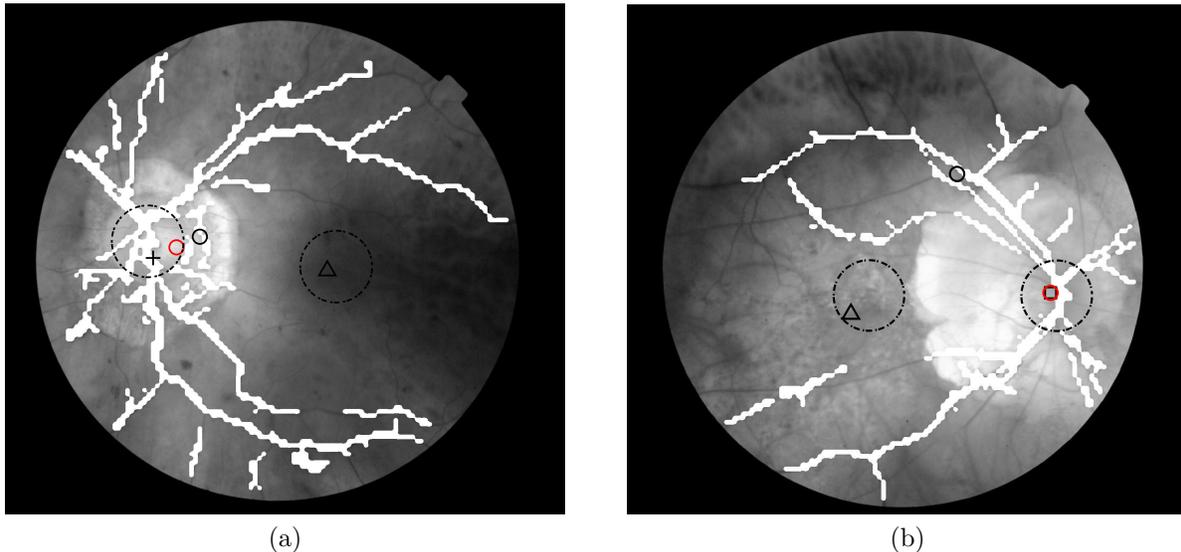


Figura 4.3: Detección correcta del DO tras considerar cuasi-equivalentes las posiciones de DO y HV (ver sección 4.3.3). Las imágenes corresponden a casos con graves patologías y en las que el método de detección original obtiene un fallo en la detección del DO (círculo negro): (a) el DO se obtiene como punto medio del DO y HV de una salida basada en tripleta ganadora; (b) el DO se obtiene como resultado de elegir el HV de una solución basada en pareja ganadora (HV, mácula) (En ambos casos la nueva solución de DO se representa mediante un círculo rojo).

igualando el mejor resultado obtenido en la literatura de detección de DO en MESSIDOR (ver tabla 4.1). De un lado, se está mejorando nuestro porcentaje de detección del DO en un 0.42% y, de otro lado, se refuerza la idea de que el HV puede utilizarse como información redundante para la detección del DO. Aunque un aumento del 0.42% (5 imágenes) en el éxito de detección puede parecer no muy significativo, hay que tener en cuenta que este se produce en imágenes que son extremadamente difíciles de etiquetar por presentar severas patologías de DO que dificultan su detección. Por ejemplo, la Fig. 4.3 muestra dos de las cinco imágenes mencionadas y en las que se puede apreciar severos signos de patología.

#### 4.4. Diseño de experimentos de segmentación

De forma similar a como se hizo en la sección 4.1 para detección, en esta sección se describen los diferentes experimentos llevados a cabo para evaluar el método de segmentación.

Además, se detallan los valores asociados a la configuración de parámetros elegida para el método, así como los criterios utilizados para evaluar los resultados de segmentación obtenidos.

Se ha diseñado una serie de experimentos para evaluar diferentes aspectos del método de segmentación. Gran parte de ellos se evalúan sobre la base de imágenes MESSIDOR y, el resto, sobre ONHSD, otra conocida base de imágenes también pública. No obstante, hay que decir que son pocos los trabajos que hemos encontrado donde se realice la segmentación simultánea de estructuras retinianas. Gran parte de los trabajos existente en la literatura realizan la segmentación aislada de una única estructura. Así, el **primer experimento** consiste en analizar si el uso conjunto del CRintra e inter-e mejora los resultados de segmentación respecto a cuando solo se usa CRintra-e. Para el primer caso, se utiliza el método propuesto en este trabajo y que, de aquí en adelante, denominaremos  $SegSim_{-}(DO+fóv)$  (segmentación simultánea de DO y fóvea), para distinguirlo de otras variantes aquí implementadas. Para el segundo caso, se implementaron dos nuevas variantes del método de segmentación basadas solo en los respectivos modelos intra-estructura ya definidos para DO (ver sección 3.2.2) y fóvea (ver sección 3.2.3). En un **segundo experimento**, se analiza y compara el efecto de realizar la segmentación de estructuras anatómicas de una forma secuencial frente a la alternativa, aquí propuesta, de hacerla de forma simultánea ( $SegSim_{-}(DO+fóv)$ ). En la segmentación secuencial de dos estructuras, se asume que, para segmentar la primera solo se utiliza CRintra-e y, para la segunda, ambos tipos de conocimiento. A su vez, es posible realizar primero la segmentación de la fóvea y utilizar el resultado de esta para segmentar el DO o a la inversa. Al método que implementa el primer tipo de segmentación se denomina  $SegSec_{-}(fóv \rightarrow DO)$  y, al segundo,  $SegSec_{-}(DO \rightarrow fóv)$ . Por el contrario, en la búsqueda simultánea de DO y fóvea, ambos tipos de conocimiento (CRintra e inter-e), son utilizados en la búsqueda de las dos estructuras. En este caso, el método que lo implementa es el ya propuesto originalmente en este trabajo ( $SegSim_{-}(DO+fóv)$ ). En el **tercer experimento**, se realiza una comparativa de los resultados de segmentación de nuestro método  $SegSim_{-}(DO+fóv)$ , con aquellos métodos en la literatura que han hecho segmentación simultánea de DO y fóvea. El **cuarto experimento** va dirigido a comparar el método  $SegSim_{-}(DO+fóv)$  con otros

métodos competitivos en la literatura dedicados a abordar el problema de segmentación de una única estructura. En este caso se realiza una comparativa con métodos que solo abordan la segmentación de DO. El **quinto experimento** es similar al anterior, pero la comparativa se hace con métodos que solo abordan la segmentación de fóvea.

Finalmente, y a fin de añadir evidencia sobre la robustez del método de segmentación presentado, este se evalúa en la base de imágenes ONHSD. Aunque existen otras bases de imágenes públicas en la literatura, la elección de ONHSD obedece a dos fuertes limitaciones: (i) no tenemos conocimiento de trabajos en la literatura donde se aborde la segmentación simultánea en esas otras bases de imágenes; y (ii) no disponemos del ground truth de contorno de DO y de fóvea en dichas bases de imágenes.

De otro lado, con el fin de evaluar los resultados de la segmentación, se han utilizado diferentes índices ya utilizados en la literatura relacionada. Por ejemplo, en la segmentación del DO, los *coeficientes de Jaccard's* (JC) y *Dice's* (D) son frecuentemente usados (ver Ecs. (4.2) y (4.3)), los cuales describen el grado de superposición entre el área *gold standard*,  $A_{GS}$ , y el área segmentada,  $A_S$ , siendo igual a 1 cuando la segmentación es perfecta. Otros índices típicamente utilizados son la *precisión*, la *fracción de verdaderos positivos* y la *fracción de falsos positivos*. La precisión ( $A_c$ ) se determina por la suma de píxeles correctamente clasificados como DO y no-DO dividida por el número total de píxeles en la imagen. La fracción de verdaderos positivos, TPF, del inglés *True Positive Fraction*, se obtiene dividiendo los píxeles clasificados correctamente como DO entre el número total de píxeles DO en el *gold standard*. La fracción de falsos positivos, FPF, del inglés *False Positive Fraction*, se calcula dividiendo los píxeles mal clasificados como DO entre el número total de píxeles no-DO en el *gold standard*.

$$JC = \frac{|A_{GS} \cap A_S|}{|A_{GS} \cup A_S|} \quad (4.2)$$

$$D = \frac{2|A_{GS} \cap A_S|}{|A_{GS}| + |A_S|} \quad (4.3)$$

Por otro lado, en relación con la segmentación de fóvea, se utilizan diferentes criterios

denotados por  $(\frac{1}{8})R_{DO}$ ,  $(\frac{1}{4})R_{DO}$ ,  $(\frac{1}{2})R_{DO}$ ,  $1R_{DO}$  y  $2R_{DO}$ . Estos criterios están basados en el porcentaje de casos donde la distancia entre la fovea estimada y el centro real de la fovea son menores que  $\frac{1}{8}$ ,  $\frac{1}{4}$ ,  $\frac{1}{2}$ , 1 y  $2R_{DO}$ , respectivamente, siendo  $R_{DO}$  calculado a partir de la Ec. (3.2). Dado que, aquí, estamos interesados en obtener el centro de fovea con la mayor precisión posible, será de especial interés los resultados obtenidos para un umbral igual o menor a  $(\frac{1}{2})R_{DO}$ , tal y como se estableció en la sección 1.2.1.

Además, se ha utilizado también la sensibilidad (Sens) y la especificidad (Espec) para evaluar la segmentación tanto de DO como de fovea. La primera de ellas es equivalente al TPF y la segunda se define como la fracción de verdaderos negativos TNF (del inglés *True Negative Fraction*) siendo equivalente a 1-FPF. Aquí es preciso mencionar que, para calcular los valores de sensibilidad y especificidad en fovea, se utilizaron los círculos centrados en el centro real de fovea y en el centro estimado de fovea como el *gold standard* de fovea y la región estimada de fovea, respectivamente. El diámetro de ambos círculos fue el mismo e igual al tamaño de subventana elegido para construir el modelo de fovea ( $L = 27 px$ , a resolución  $K_0$ ).

En cuanto a la configuración del método propuesto, la Tabla 4.7 muestra el valor usado para los diferentes parámetros del algoritmo DE. Estos valores fueron obtenidos experimentalmente, sin necesidad de tener que realizar un ajuste minucioso para obtener buenos resultados.

Finalmente, y dado el carácter estocástico de los algoritmos evolutivos, todos los experimentos de esta sección se han repetido 10 veces. Por consiguiente, los resultados proporcionados en las tablas relacionadas corresponden a los valores medios de los resultados para las 10 ejecuciones, junto con sus respectivas desviaciones típicas. Cabe indicar que, a fin de no sobrecargar las tablas, solamente se han indicado dichas desviaciones típicas en las comparativas de aquellos trabajos que también lo hacen.

Tabla 4.7: Valor de parámetros utilizados en el algoritmo DE

Parámetro	Valor
Tipo de estrategia DE	“DE/rand/1/bin”
Tamaño de Población	25
Criterio de finalización	#Generaciones = 800
Probabilidad Cruce ( $CR$ )	0.7
Peso diferencial ( $F$ )	0.8
Función <i>fitness</i>	Ver Ec. (3.29)
$[\omega_{DO}, \omega_F]$	[0.5, 0.5]

## 4.5. Resultados de segmentación

Los resultados de segmentación se presentan de forma paralela a los experimentos definidos en la sección anterior.

### 4.5.1. Evaluación en la base de imágenes MESSIDOR

En esta sección se describen todos los resultados obtenidos con aquellos experimentos definidos y evaluados en MESSIDOR. Siguiendo la definición del primer experimento, se implementaron dos nuevas variantes del método original de segmentación, denominadas  $Seg_{-}(DO)$  y  $Seg_{-}(fóv)$ , con el fin de poder comparar el método original, que usa ambos tipos de conocimiento, con sendas variantes que solo usan CRintra-e.

De un lado, el método  $Seg_{-}(DO)$  realiza una búsqueda evolutiva del contorno de DO usando solo CRintra-DO. En este caso, al prescindir del conocimiento relacional de fóvea, cada individuo de la población se codifica con solo cinco dimensiones ( $\widehat{C}_{DOx}, \widehat{C}_{DOy}, \widehat{a}, \widehat{b}, \widehat{\rho}$ ). Para evaluar cada individuo, se aplica la función dada por la Ec. (3.25), teniendo en cuenta que  $\widehat{M}_{DO}$  se calcula a partir del Algoritmo 3.2 e incorporando los cambios mostrados en el algoritmo 4.1. Los cambios obedecen a las siguientes motivos: (1) si solo se quiere hacer uso del CRintra-DO, habrá que prescindir de la información de vasos utilizada en la búsqueda

## 4.5. RESULTADOS DE SEGMENTACIÓN

**Algoritmo 4.1** Los tres cambios a realizar en el pseudocódigo del algoritmo 3.2 para poder implementar el método  $Seg_-(DO)$ .

---

```

(1) %Eliminar línea 2
(2) %Modificar línea 4 por:
    [pDO, paux] := CalculaPtosCorteHoriz(eDO, eaux);
(3) %Modificar líneas 7-18 por:
    if ING( $\tilde{\mathbf{e}}_{DO}(j)$ ) > ING( $\tilde{\mathbf{e}}_{aux}(k)$ ),  $\forall k \in \{j - \frac{n}{2}, \dots, j, \dots, j + \frac{n}{2}\}$ 
         $\widehat{\mathbf{M}}_{DO}(j) := \mathbf{M}_{DO}(j)$ ;
    else
         $\widehat{\mathbf{M}}_{DO}(j) := 0$ ;
    end-if

```

---

Tabla 4.8: Influencia de usar o no el conocimiento relacional inter-estructura (CRinter-e) en el proceso de segmentación: los métodos  $Seg_-(fóv)$  y  $Seg_-(DO)$  solo usan CRintra-e de fóvea y DO, respectivamente. El método  $SegSim_-(DO+fóv)$  usa ambos tipos de conocimiento.

Método	Fóvea (%)					DO				
	$\frac{1}{8}R_{DO}$	$\frac{1}{4}R_{DO}$	$\frac{1}{2}R_{DO}$	$1R_{DO}$	$2R_{DO}$	JC	D	Ac	TPF	FPF
Seg <sub>-</sub> (DO)	-	-	-	-	-	0.8418	0.8983	0.9981	0.8730	0.0005
Seg <sub>-</sub> (fóv)	74.1	91.1	93.6	94.3	94.6	-	-	-	-	-
SegSim <sub>-</sub> (DO+fóv)	<b>74.7</b>	<b>94.1</b>	<b>97.3</b>	<b>99.0</b>	<b>99.6</b>	<b>0.8730</b>	<b>0.9252</b>	<b>0.9985</b>	<b>0.9072</b>	<b>0.0005</b>

del DO; (2) por la misma razón que antes, habrá que prescindir del CRintra-e de fóvea ( $\mathbf{C}_F$ ) durante la búsqueda. Por tanto, la denominada función `CalculaPtosCorteHoriz` calculará ahora los puntos de corte de las elipses genética y auxiliar con la recta horizontal que atraviesa el centro de la elipse; y (3) al prescindir de la información de vasos y de fóvea, ya solo se utiliza una heurística basada en CRintra-DO: el contorno papilar separa el interior del DO (zona brillante) del resto de la retina (menos brillante).

De otro lado, el método  $Seg_-(fóv)$  realiza una búsqueda evolutiva de la fóvea utilizando solamente CRintra-fóvea. En este caso, al prescindir del conocimiento relacional de DO, cada individuo de la población se codifica con dos dimensiones ( $\widehat{\mathbf{C}}_{Fx}, \widehat{\mathbf{C}}_{Fy}$ ), que corresponden al centro,  $\widehat{\mathbf{C}}_F$ , de la subventana  $\widehat{\mathbf{W}}_F$ . Para evaluar cada individuo, se aplica la función dada por la Ec. (3.26), teniendo en cuenta que  $\widehat{\mathbf{M}}_F$  se calcula mediante el Algoritmo 3.4. La Tabla 4.8 muestra la comparativa de resultados obtenidos por los tres métodos.

En el segundo experimento, se analiza y compara el efecto de realizar la segmentación de estructuras anatómicas de una forma secuencial frente a la alternativa, del método original de segmentación, de hacerla de forma simultánea. Para ello se implementan dos nuevas variantes del método original, denominadas  $SegSec_{-}(fóv \rightarrow DO)$  y  $SegSec_{-}(DO \rightarrow fóv)$ , que realizan la segmentación secuencial de fóvea y DO en el orden indicado para el primer caso y, de DO y fóvea, en el segundo. Así, el método  $SegSec_{-}(fóv \rightarrow DO)$  consta de dos etapas que se describen a continuación. En la primera etapa, se realiza una búsqueda evolutiva de la fóvea usando solo CRintra-fóvea, es decir, mediante el método  $Seg_{-}(fóv)$  definido anteriormente. Seguidamente, se realiza una búsqueda evolutiva del DO basada en el uso de CRintra e inter-e, incluyendo el resultado de fóvea obtenido en la etapa anterior. En este caso, cada individuo de la población se codifica mediante cinco parámetros:  $(\hat{a}, \hat{b}, \hat{\rho}, \hat{\varphi}, \hat{\varepsilon})$ . Conocidos los valores de  $(\hat{\varphi}, \hat{\varepsilon})$  y de  $(\hat{C}_{Fx}, \hat{C}_{Fy})$ , estos dos últimos obtenidos como resultado de aplicar la primera etapa, es posible utilizar la Ec. (3.21) para calcular el centro de DO  $(\hat{C}_{DOx}, \hat{C}_{DOy})$ . Finalmente, conocidos  $(\hat{C}_{DO}, \hat{a}, \hat{b}, \hat{\rho}, \hat{C}_F)$ , es posible evaluar cada individuo mediante la función dada por la Ec. (3.25), y teniendo en cuenta que  $\hat{M}_{DO}$  se calcula a partir del Algoritmo 3.2.

De otro lado, las dos etapas del método  $SegSec_{-}(DO \rightarrow fóv)$  se implementan como se indica a continuación. En la primera etapa, se realiza una búsqueda evolutiva del DO usando solo CRintra-DO, es decir, mediante el método  $Seg_{-}(DO)$  definido anteriormente. Seguidamente, se realiza una búsqueda evolutiva de la fóvea basada en el uso conjunto de CRintra e inter-e, incluyendo el resultado de DO obtenido en la etapa anterior. En este caso, cada individuo de la población se codificará con solo dos parámetros:  $(\hat{\varphi}, \hat{\varepsilon})$ . Conocido el valor de estos dos parámetros y de  $(\hat{C}_{DOx}, \hat{C}_{DOy})$ , estos dos últimos obtenidos como resultado de aplicar la primera etapa, se puede usar la Ec. (3.21) para obtener el centro de fóvea  $(\hat{C}_{Fx}, \hat{C}_{Fy})$ . Finalmente, conocido  $\hat{C}_F$  es posible evaluar cada individuo mediante la función dada por la Ec. 3.26, y teniendo en cuenta que  $\hat{M}_F$  se calcula mediante el Algoritmo 3.4. La Tabla (4.9) muestra una comparativa de los dos métodos de segmentación secuencial respecto al método de segmentación simultánea propuesto en este trabajo.

En el tercer experimento se realiza una comparativa de los resultados de segmentación de nuestro método con aquellos métodos en la literatura que han hecho segmentación simultánea

## 4.5. RESULTADOS DE SEGMENTACIÓN

Tabla 4.9: Comparativa de realizar segmentación de estructuras de forma simultánea ( $SegSim_{-}(DO+fóv)$ ) o secuencial ( $SegSec_{-}(fóv \rightarrow DO)$  y  $SegSec_{-}(DO \rightarrow fóv)$ ).

Método	Fóvea (%)					DO				
	$\frac{1}{8}R_{DO}$	$\frac{1}{4}R_{DO}$	$\frac{1}{2}R_{DO}$	$1R_{DO}$	$2R_{DO}$	JC	D	Ac	TPF	FPF
SegSec <sub>-</sub> (fóv→DO)	74.2	91.1	93.6	94.3	94.6	0.8309	0.8803	0.9977	0.8635	0.0009
SegSec <sub>-</sub> (DO→fóv)	<b>76.4</b>	94.0	96.9	98.8	99.2	0.8418	0.8983	0.9981	0.8730	0.0005
SegSim <sub>-</sub> (DO+fóv)	74.7	<b>94.1</b>	<b>97.3</b>	<b>99.0</b>	<b>99.6</b>	<b>0.8730</b>	<b>0.9252</b>	<b>0.9985</b>	<b>0.9072</b>	<b>0.0005</b>

Tabla 4.10: Comparativa con otros métodos de segmentación simultánea de DO y fóvea.

Método	Base de imág. (#)	Fóvea				DO		
		Sens	Espec	$\leq \frac{1}{2}R_{DO}$	$\leq 1R_{DO}$	Sens	Espec	JC
[Girard et al., 2016]	MESSIDOR (1200)	-	-	94.0%	98.0%	-	-	<b>0.90</b>
[Tan et al., 2017]	DRIVE (40)	<b>0.8853</b>	0.9914	-	-	0.8790	0.9927	0.62
<b>Nuestro método</b>	MESSIDOR (1200)	0.8841	<b>0.9996</b>	<b>97.3%</b>	<b>99.0%</b>	<b>0.9072</b>	<b>0.9995</b>	0.87

de DO y fóvea. Los resultados de dicha comparativa se muestran en la Tabla 4.10.

Los experimentos cuarto y quinto, realizan sendas comparativas del método de segmentación simultánea con otros métodos competitivos en la literatura dedicados a abordar el problema de segmentación de una única estructura. Así, mientras que la Tabla 4.11 muestra una comparativa con métodos que solo abordan la segmentación de DO, la Tabla 4.12 hace lo propio con métodos que solo abordan la segmentación de fóvea.

Finalmente, desde el punto de vista de la información de salida producida por el método de segmentación simultánea propuesto en este trabajo, la Fig. 4.4 muestra algunos ejemplos de segmentación obtenidos para diferentes imágenes MESSIDOR. Se ha detectado que el método puede fallar en aquellas imágenes donde existen signos severos de patologías relacionadas

Tabla 4.11: Comparativa con otros métodos competitivos de la literatura que solo abordan el problema de segmentación de DO (la etiqueta DE denota la Desviación Estándar). En todos los casos, los métodos fueron evaluados en las 1200 imágenes MESSIDOR.

Método	JC $\pm$ DE	D $\pm$ DE	Ac $\pm$ DE	TPF $\pm$ DE	FPF $\pm$ DE
[Aquino et al., 2010]	0.8600	-	-	-	-
[Morales et al., 2013]	0.8228 $\pm$ 0.1384	0.8950 $\pm$ 0.1056	0.9949 $\pm$ 0.0050	0.9300 $\pm$ 0.1239	0.0035 $\pm$ 0.0041
[Dashtbozorg et al., 2015]	<b>0.8859 <math>\pm</math></b> <b>0.0818</b>	<b>0.9373 <math>\pm</math></b> <b>0.0509</b>	<b>0.9987 <math>\pm</math></b> <b>0.0012</b>	<b>0.9481 <math>\pm</math></b> <b>0.0459</b>	0.0008 $\pm$ 0.0010
<b>Nuestro método</b>	0.8730 $\pm$ 0.1254	0.9252 $\pm$ 0.1068	0.9985 $\pm$ 0.0020	0.9072 $\pm$ 0.1224	<b>0.0005 <math>\pm</math></b> <b>0.0008</b>

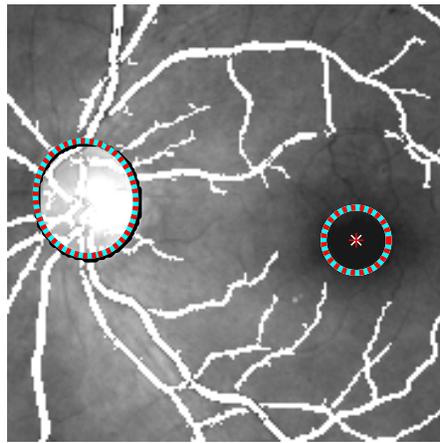
Tabla 4.12: Comparativa (en tanto por ciento) con otros métodos competitivos de la literatura que solo abordan el problema de segmentación de fovea. En todos los casos, los métodos fueron evaluados en las 1200 imágenes MESSIDOR.

Método	$\frac{1}{8}R_{DO}$	$\frac{1}{4}R_{DO}$	$\frac{1}{2}R_{DO}$	$1R_{DO}$	$2R_{DO}$
[Gegundez et al., 2013]	<b>80.4</b>	93.9	96.1	96.9	97.8
[Kao et al., 2014]	-	-	<b>97.8</b>	98.2	-
<b>Nuestro método</b>	74.7	<b>94.1</b>	97.3	<b>99.0</b>	<b>99.6</b>

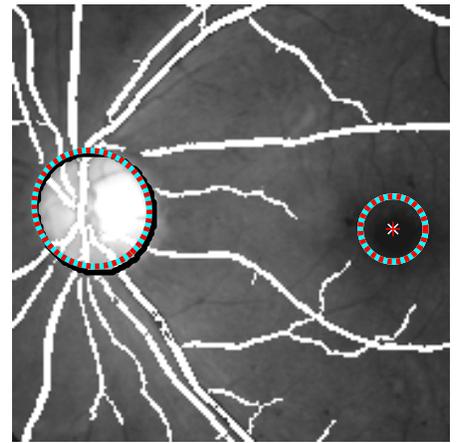
tanto con la mácula/fóvea (alto número de microaneurismas, hemorragias o grandes exudados duros en el área macular) como con el DO (atrofias peripapilares avanzadas o estafilomas posteriores), tal y como se muestra, por ejemplo, en las subfiguras (e) y (f) de la Fig. 4.4, respectivamente.

#### 4.5.2. Evaluación en la base de imágenes ONHSD

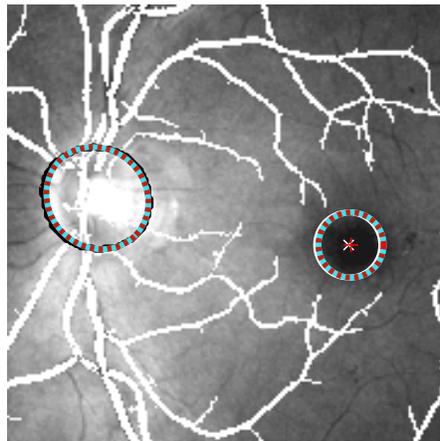
En esta sección se muestran los resultados de aplicar el método de segmentación simultánea la base de imágenes ONHSD. Sin embargo, cabe indicar que, en la bibliografía manejada, no se encontraron trabajos que segmentaran la fovea para esta base de imágenes con un criterio de distancia igual o menor que  $\frac{1}{2}R_{DO}$ . Además, aquí también se hace uso de un índice



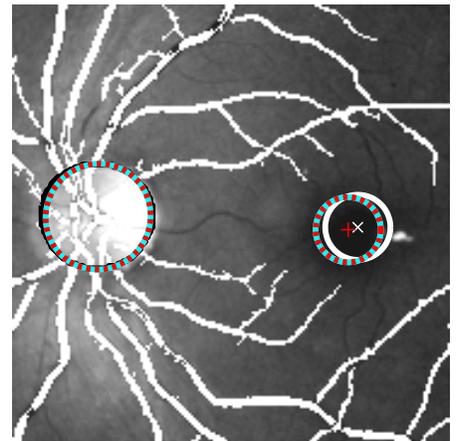
(a)



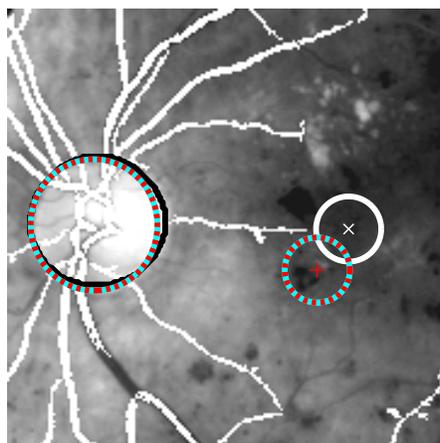
(b)



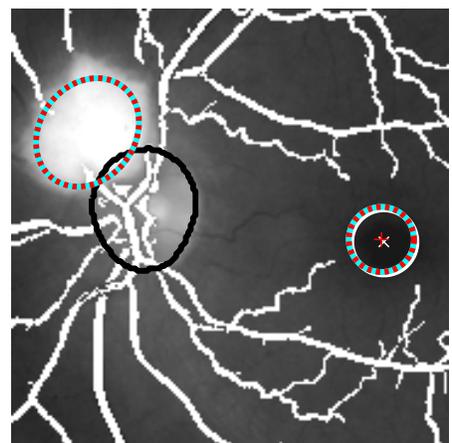
(c)



(d)



(e)



(f)

Figura 4.4: Ejemplos de segmentación obtenidos para diferentes imágenes MESSIDOR: ambas estructuras son segmentadas con muy buena aproximación (a-b) o con una buena aproximación (c-d); solo es correctamente segmentado el DO (e) o la fovea (f). La línea continua representa el *gold standard* y, la línea a trazos, la segmentación obtenida.

de medida, denominado *discrepancia* [Lowell et al., 2004], que ha sido utilizado ampliamente en la literatura para evaluar los resultados de segmentación del DO en la mencionada base de imágenes. La discrepancia  $\delta$  es una medida de distancia entre el contorno obtenido y el contorno *gold standard* cuando este último se obtiene como el promedio de los contornos trazados por dos o más expertos. Concretamente, la discrepancia  $\delta^j$ , dada la imagen  $j$ -ésima, se define como:

$$\delta^j = \sum_{i=1}^P \frac{|m_i^j - \mu_i^j|}{\sigma_i^j + \epsilon} \quad (4.4)$$

siendo  $m_i^j$  la longitud del radio del  $i$ -ésimo punto del contorno segmentado para la imagen  $j$ -ésima, medido respecto de un origen de coordenadas situado en el centro real del DO. Por otra parte,  $\mu_i^j$  y  $\sigma_i^j$  son la media y la desviación típica, respectivamente, de las longitudes de los radios definidas para el  $i$ -ésimo punto de los contornos trazados por los expertos y pertenecientes a la imagen  $j$ -ésima. Además,  $\epsilon \rightarrow 0$  es una constante que evita la división por cero cuando no hay discrepancia entre los expertos en un punto del contorno dado. Finalmente, se definen también un conjunto de etiquetas que dan una medida subjetiva de la calidad de la segmentación resultante [Lowell et al., 2004]: Excelente ( $\delta < 1$ ), Buena ( $\delta \in [1, 2]$ ), Aceptable ( $\delta \in [2, 5]$ ) y Pobre ( $\delta > 5$ ). Nótese que la medida de discrepancia no puede ser usada en MESSIDOR debido a que solo se dispone del trazado de contorno del DO de un experto por cada imagen. La Tabla 4.13 muestra los resultados de segmentación obtenidos por nuestro método comparados con otros métodos de segmentación de DO competitivos en la literatura para la base de imágenes ONHSD. Finalmente, la tabla 4.14 muestra los resultados de segmentación obtenidos por nuestro método para fóvea.

## 4.5. RESULTADOS DE SEGMENTACIÓN

Tabla 4.13: Comparativa con métodos competitivos de la literatura que solo abordan el problema de segmentación de DO (la etiqueta DE denota la Desviación Estándar). En todos los casos, los métodos fueron evaluados en las 87 imágenes ONHSD.

Método	JC ± DE	D ± DE	Ac ± DE	TPF ± DE	FPF ± DE	$\delta$ Excel	$\delta$ Buena	$\delta$ Acept	$\delta$ Pobre
[Aquino et al., 2010]	-	-	-	-	-	40 %	39 %	18 %	3 %
[Morales et al., 2013]	0.8045 ± 0.1175	0.8867 ± 0.0776	0.9941 ± 0.0042	0.9310 ± 0.1046	0.0043 ± 0.0042	28 %	36 %	<b>31 %</b>	6 %
[Girard et al., 2016]	0.8400	-	-	-	-	<b>51 %</b>	36 %	13 %	<b>0 %</b>
[Dashtbozorg et al., 2015]	0.8341 ± 0.0912	<b>0.9173 ±</b> <b>0.0634</b>	0.9968 ± 0.0027	<b>0.9435 ±</b> <b>0.0791</b>	0.0012 ± 0.0027	44 %	47 %	9 %	<b>0 %</b>
<b>Nuestro método</b>	<b>0.8417 ±</b> <b>0.1132</b>	0.9083 ± 0.0971	<b>0.9979 ±</b> <b>0.0022</b>	0.8756 ± 0.1153	<b>0.0005 ±</b> <b>0.0010</b>	25 %	<b>57 %</b>	15 %	3 %

Tabla 4.14: Resultados (en tanto por ciento) de segmentación de fovea para la base de imágenes ONHSD (87 imágenes).

Método	$\frac{1}{8}R_{DO}$	$\frac{1}{4}R_{DO}$	$\frac{1}{2}R_{DO}$	$1R_{DO}$	$2R_{DO}$
<b>Nuestro método</b>	69.43	93.68	96.44	98.85	99.89

## 4.6. Discusión de resultados de segmentación

A continuación, se realiza una discusión de los resultados obtenidos en la sección anterior.

### 4.6.1. Base de imágenes MESSIDOR

En relación a los dos primeros experimentos que analizan la influencia del CRinter-e y la ventaja o no de realizar la segmentación de forma simultánea, se puede afirmar lo siguiente: (i) los resultados del primer experimento (ver Tabla 4.8) muestran que es ventajoso el uso conjunto de CRintra e inter-e frente al uso únicamente de CRintra-e; y (ii) a la vista de los resultados obtenidos por el segundo experimento (ver Tabla 4.9), se puede apreciar la ventaja de usar segmentación simultánea frente a la secuencial. Por tanto, los resultados de estos dos primeros experimentos refuerzan las dos primeras hipótesis propuestas en esta tesis.

El método de segmentación simultánea también ha demostrado ser competitivo en comparación con otros trabajos de segmentación simultánea existentes en la literatura (ver Tabla 4.10). En relación al DO, nuestro método es solo ligeramente superado por [Girard et al., 2016] (siendo mejores nuestros resultados de segmentación de fóvea). En relación a la fóvea, nuestro método es solo ligeramente sobrepasado por [Tan et al., 2017] (siendo mejores nuestros resultados de segmentación de DO). Sin embargo, es preciso señalar que este último método fue evaluado en un base de imágenes (DRIVE) de solo 40 imágenes.

Si comparamos los resultados de segmentación simultánea con aquellos trabajos que segmentan DO y fóvea de forma aislada, tenemos que nuestros resultados también son competitivos. En relación a la comparativa con métodos que solo segmentan el DO de forma aislada (ver Tabla 4.11), nuestros resultados son solo superados ligeramente por [Dashtbozorg et al., 2015]. No obstante, un análisis más detenido de estos resultados revela que los valores obtenidos por nuestro método para TPF y FPF son los más bajos de los presentados en la comparativa, indicando con ello que es el método más conservativo, es decir, identifica un menor número de píxeles reales de DO (verdaderos positivos), pero con la ventaja de detectar un menor número de falsos positivos. Este comportamiento podría estar motivado por la forma elíptica elegida por nuestro modelo para aproximar el contorno del DO. Así, en

[Dashtbozorg et al., 2015], un valor ligeramente superior para TPF y un valor parecido para FPF, comparados con los obtenidos en nuestro caso, pueden explicarse porque la búsqueda del contorno de DO allí es más flexible que en nuestro método, dado que no tiene que adaptarse a ninguna forma geométrica regular previa. Por el contrario, en [Morales et al., 2013], un mayor valor para TPF, pero a expensas de un elevado valor de FPF, pueden deberse a que la búsqueda allí es aún más rígida que en nuestro modelo, al aproximar el contorno de DO mediante una circunferencia. Por otra parte, y en relación a la comparativa con métodos que segmentan de forma aislada la fóvea (ver Tabla 4.12), nuestro método sigue siendo competitivo. Además, hay que tener en cuenta que, en [Gegundez et al., 2013], el método de segmentación de fóvea implementado requiere, como parámetros de entrada, el diámetro y las coordenadas del centro del DO, además de una segmentación de los vasos de la imagen a analizar. En particular, dichas estructuras se obtienen previamente tras aplicar la segmentación del DO [Aquino et al., 2010] y de los vasos [Marin et al., 2011]. De otro lado, en [Kao et al., 2014], la segmentación de fóvea requiere también de dos etapas previas: una etapa dedicada a realizar la detección del centro de DO y otra dedicada a determinar la región libre de vasos en la que ubicar el eje DO-fóvea. Es decir, en ambos casos se realiza una segmentación de fóvea, siguiendo un esquema de segmentación secuencial de estructuras, en el que previamente el DO es segmentado o detectado.

Por último, es importante señalar que la búsqueda evolutiva de nuestro método puede desatornillarse del modelo relacional propuesto y, por consiguiente, del grado de aproximación del modelo intra-e construido para cada estructura. Esto significa que cuanto mejor sea cada uno de los modelos intra-e usados, mejores resultados se deberían obtener en la búsqueda evolutiva de la instancia del modelo relacional resultante.

### 4.6.2. Base de imágenes ONHSD

Con respecto a la base de imágenes ONHSD, los resultados de segmentación obtenidos para DO son también competitivos (ver Tabla 4.13). Así, se obtienen mejores resultados que en [Morales et al., 2013, Girard et al., 2016] y resultados similares a [Dashtbozorg et al., 2015] para los índices numéricos (JC, D, Ac, TPF y FPF). En cuanto a las medidas de discrepancia,

nuestro método obtiene el mejor resultado en la categoría *Bueno*, siendo el más bajo en la categoría de *Excelente*. Este último resultado puede ser debido a la rigidez del modelo elíptico impuesto para el contorno. En [Aquino et al., 2010, Morales et al., 2013] ocurre algo similar, ya que, en ambos casos, el contorno devuelto por el método es aproximado siempre por un círculo, si bien es cierto que, en [Aquino et al., 2010], el porcentaje de éxito en la categoría *Excelente* no se ve muy afectado. Al contrario, en [Dashtbozorg et al., 2015, Girard et al., 2016] se trabaja con contornos con forma no predeterminada, lo que posibilita una aproximación más exacta al contorno real, y por consiguiente, mejores medidas de discrepancia en la categoría más restrictiva. Además, hay que recordar que nuestro método está segmentando dos estructuras simultáneamente y no es un método optimizado para centrarse en una única estructura. Los resultados de segmentación de fovea (ver Tabla 4.14) no pueden ser comparados ya que no conocemos ningún trabajo que haya usado esta base de imágenes para localizar la fovea con un criterio de distancia menor o igual que  $\frac{1}{2}R_{DO}$ . No obstante, puede observarse que se obtienen resultados muy similares a los obtenidos en MESSIDOR para umbrales  $\leq \frac{1}{2}R_{DO}$  y  $\leq \frac{1}{4}R_{DO}$ , aunque ligeramente inferiores para el umbral  $\leq \frac{1}{8}R_{DO}$ .

### 4.6.3. Coste computacional

Finalmente, se analiza el coste computacional del método de segmentación. El método fue implementado usando lenguaje MATLAB y en un portátil con CPU Intel Core i7-2640M, con 2.8GHz y 8GB de RAM. La tabla 4.15 muestra una comparativa con otros métodos en la literatura que hacen segmentación simultánea de DO y fovea. Dado que estos algoritmos fueron implementados en diferentes equipos informáticos, el tiempo de ejecución debería interpretarse en función del hardware, del lenguaje de programación utilizados y de los resultados de segmentación obtenidos (véase Tabla 4.10). Atendiendo a las referencias manejadas en esta tesis, el trabajo de [Tan et al., 2017] es el primer y único estudio conocido en el que se aplican redes neuronales convolucionales a la segmentación simultánea de DO y fovea en imágenes de retina. Aunque los resultados que utilizan esta técnica son prometedores, el principal inconveniente es que se necesita una enorme cantidad de muestras para entrenar y mucho más tiempo que otros algoritmos de segmentación para realizar la segmentación: se

## 4.7. RESULTADOS DE SEGMENTACIÓN CON INFORMACIÓN DE DETECCIÓN

---

Tabla 4.15: Comparativa con otros métodos de segmentación simultánea de DO y fóvea, considerando el hardware, lenguaje de programación y tiempo de computación.

Método	Hardware	Lenguaje	Tiempo (s)
[Girard et al., 2016]	Multi-core processing	C++	6.4
[Tan et al., 2017]	Two Intel Xeon, 2.66GHz, 24GB RAM	MATLAB	2035
<b>Este trabajo</b>	Intel Core i7-2640M, 2.8GHz, 8GB RAM	MATLAB	29.35

necesita 9.83 horas por época para entrenar la red, y la segmentación completa de una única imagen necesita 33.92 min (si el algoritmo se ejecuta con GPUs, *Graphics Processing Unit* en inglés, la segmentación puede emplear aproximadamente unos 10 min).

## 4.7. Resultados de segmentación con información de detección

Finalmente, se plantean aquí dos experimentos dirigidos a evaluar el método de segmentación simultánea con información de detección (ver sección 3.3). El **primer experimento** consiste en comparar dicho método, que a partir de ahora denotaremos por *Detec+Segm*, con el método de segmentación simultánea, *SegSim<sub>-(DO+fóv)</sub>*. El objetivo aquí es estudiar si el uso de los resultados de detección en la segmentación puede producir una mejora de esta última. Para realizar una comparación justa, el tamaño de la población (25 individuos) y el número de generaciones (800) se mantuvo igual en ambos métodos. El tamaño de la denominada *población temporal* (ver algoritmo 3.6), necesaria para crear la población inicial del método *Detec+Segm*, se eligió a  $N = 1000$ . La Tabla 4.16 muestra los resultados de esta comparación sobre la base de imágenes MESSIDOR. En ella puede apreciarse que, efectivamente, el conocimiento de detección produce una mejora en los resultados de segmentación. El pequeño inconveniente en el método *Detec+Segm* es el incremento de tiempo ( $\Delta T$ ) asociado al coste de la etapa de detección y de generación de la población inicial. Dicho incremento se puede cuantificar de la siguiente forma:

Tabla 4.16: Influencia de usar o no el conocimiento relacional inter-estructura (CRinter-e) en el proceso de segmentación:  $SegSim_{-}(DO+fóv)$  es el método de segmentación simultánea y, Detec+Segm, el método de segmentación simultánea con información de detección (Resultados para MESSIDOR).

Método	Fóvea (%)					DO				
	$\frac{1}{8}R_{DO}$	$\frac{1}{4}R_{DO}$	$\frac{1}{2}R_{DO}$	$1R_{DO}$	$2R_{DO}$	JC	D	Ac	TPF	FPF
Detec+Segm	<b>75.26</b>	<b>94.22</b>	<b>97.50</b>	<b>99.28</b>	<b>99.80</b>	<b>0.8804</b>	<b>0.9314</b>	<b>0.9986</b>	<b>0.9144</b>	<b>0.0004</b>
SegSim <sub>-</sub> (DO+fóv)	74.71	94.11	97.29	99.04	99.58	0.8730	0.9252	0.9985	0.9072	0.0005

$$\Delta T = \frac{(T_D + T_S) - T_{S_0}}{T_{S_0}} \times 100 \quad (4.5)$$

Siendo  $T_D = 4.17s$  y  $T_{S_0} = 29.35s$  los tiempos medios empleados por el método de detección y el método de segmentación original, respectivamente y  $T_S = 33.31s$  el tiempo medio que emplea el método  $Detec+Segm$ . De esta forma, el incremento de tiempo que se obtiene en relación al empleado por el método de segmentación es  $\Delta T = 27.70\%$ .

Por otra parte, el **segundo experimento** es similar al anterior, pero aplicado a la base de imágenes ONHSD. El objetivo aquí es comprobar si la mejora observada con el método  $Detec+Segm$  en MESSIDOR se mantiene también en otra base de imágenes diferente. La Tabla 4.17 muestra los resultados. Al igual que ocurría con MESSIDOR, los resultados obtenidos por el método  $Detec+Segm$  mejoran, para todos los índices, respecto de los obtenidos por el método  $SegSim_{-}(DO+fóv)$ .

Finalmente, se puede afirmar que los resultados obtenidos en estos dos experimentos apoyan y dan robustez a la tercera hipótesis planteada en esta tesis sobre el beneficio de inyectar el resultado de detección en el método de segmentación.

## 4.7. RESULTADOS DE SEGMENTACIÓN CON INFORMACIÓN DE DETECCIÓN

---

Tabla 4.17: Influencia de usar o no el conocimiento relacional inter-estructura (CRinter-e) en el proceso de segmentación: *SegSim<sub>-</sub>(DO+fóv)* es el método de segmentación simultánea y, *Detec+Segm*, el método de segmentación simultánea con información de detección (Resultados para ONHSD).

Método	JC ± DE	D ± DE	Ac ± DE	TPF ± DE	FPF ± DE	Excel	Buena	Accept	Pobre
Detec+Segm	<b>0.8487 ± 0.0863</b>	<b>0.9154 ± 0.0583</b>	<b>0.9980 ± 0.0015</b>	<b>0.8833 ± 0.0859</b>	<b>0.0005 ± 0.0007</b>	<b>26 %</b>	<b>57 %</b>	<b>16 %</b>	<b>1 %</b>
SegSim <sub>-</sub> (DO+fóv)	0.8417 ± 0.1132	0.9083 ± 0.0971	0.9979 ± 0.0022	0.8756 ± 0.1153	0.0005 ± 0.0010	25 %	<b>57 %</b>	15 %	3 %

# Capítulo 5

## Conclusiones y trabajo futuro

En este capítulo se presentan las distintas conclusiones de esta tesis, así como las líneas de trabajo futuro.

### 5.1. Conclusiones

En relación a los **objetivos planteados en esta tesis**, cabe destacar las siguientes conclusiones:

- Se han diseñado diferentes modelos basados en conocimiento intra-estructura (modelos intra-e) y en conocimiento inter-estructura (modelos inter-e) de las principales estructuras retinianas (DO, mácula/fóvea, haz vascular). La complejidad de cada uno de ellos ha dependido del tipo de tarea que se quería resolver: detección o segmentación. Esto responde a los objetivos 1 y 2 planteados inicialmente en esta tesis.
- Se han presentado dos metodologías para la detección y segmentación de las principales estructuras anatómicas en imágenes de retina. Ambas metodologías son independientes del tamaño original de la imagen, plantean la detección/segmentación de estructuras de forma simultánea y utilizan dos tipos de conocimiento relacional: intra e inter-estructura. Esto responde a los objetivos 3 y 4 planteados inicialmente.
- Se ha implementado y evaluado un método basado en la mencionada metodología de

detección dirigido a la detección simultánea de DO, mácula, haz vascular y red de vasos principales. Es rápido, sencillo y sus resultados de detección son competitivos en relación con otros métodos existentes en la literatura reciente, incluyendo métodos de detección individual, secuencial o simultánea. La principal ventaja del carácter simultáneo de la detección y del uso conjunto del CRintra e inter-e fue la reducción del número de falsos positivos, comparado con otras variantes del método donde la detección se realizó de forma aislada o secuencial, o donde solo se usó CRintra-e. No se conoce ningún método en la literatura que haya abordado el reto de detectar cuatro estructuras retinianas (ya sea secuencial o simultáneamente). Manteniendo la misma configuración de parámetros, el método se aplicó a diferentes bases de imágenes públicas y, en todas ellas, las tasas de éxito de detección obtenidas fueron consistentes, evidenciando su robustez. Esto responde al objetivo 5 planteado inicialmente.

- Se ha implementado y evaluado un método basado en la mencionada metodología de segmentación dirigido a la segmentación simultánea del DO y la fóvea. Los resultados de segmentación obtenidos son competitivos en relación con otros métodos existentes en la literatura reciente, incluyendo métodos de segmentación individual, secuencial o simultánea. Nuevamente, la principal ventaja del carácter simultáneo de la segmentación y el uso conjunto de CRintra e inter-e fue la reducción del número de falsos positivos, comparada con otras variantes del método donde la segmentación se realizó de forma aislada o secuencial, o donde solo se usó CRintra-e. Por el contrario, el principal inconveniente fue el coste computacional asociado a la búsqueda poblacional evolutiva de la solución. Manteniendo la misma configuración de parámetros, el método se aplicó a dos bases de imágenes públicas y, en ambos casos, las tasas de éxito de detección obtenidas fueron consistentes, evidenciando su robustez. Esto responde al objetivo 6 planteado inicialmente.
- Se ha implementado una estrategia que permite, al método de segmentación implementado, usar eficientemente el resultado obtenido por el método de detección. Esto responde al objetivo 7 planteado inicialmente.

- Se ha implementado y evaluado un método rápido y eficiente para la detección automática del tipo de ojo (izquierdo o derecho) mostrado en una imagen de retina. Está basado en el uso del modelo intra-e del haz vascular. Este método no estaba contemplado inicialmente en los objetivos de esta tesis.

En relación a las **hipótesis planteadas en esta tesis**, cabe destacar las siguientes conclusiones:

1. Los resultados obtenidos con los métodos implementados, tanto con el de detección como con el de segmentación, apoyan la siguiente hipótesis: El uso conjunto del CR<sub>intra</sub> e inter-e en el proceso de detección o de segmentación favorece la reducción de falsos positivos en los resultados obtenidos y, consecuentemente, aumenta la tasa de éxito respecto de aquellos casos en los que solo se usa CR<sub>intra</sub>-e.
2. Los resultados obtenidos con los métodos implementados, tanto con el de detección como con el de segmentación, apoyan la siguiente hipótesis: La detección/segmentación simultánea de  $n$  estructuras a partir del uso conjunto del CR<sub>intra</sub> e inter-e favorece la elección de aquella  $n$ -tupla de estructuras que mejor explica ambos tipos de conocimiento y, por tanto, también aumenta la bondad del proceso de detección/segmentación.
3. Los resultados obtenidos con el método de segmentación con información de detección apoyan la siguiente hipótesis: la incorporación adicional del conocimiento de detección (salida del proceso de detección) al proceso de segmentación favorece la tasa de éxito. No obstante, el coste computacional aumenta en un 28 % comparado con el método de segmentación original.

## 5.2. Trabajo futuro

Como trabajo futuro, cabe citar el posible estudio de las siguientes líneas de investigación:

- Una posible línea podría ir dirigida a la creación de modelos de estructuras retinianas más precisos. Así, por ejemplo, dado que el contorno del DO no obedece estrictamente

a una forma geométrica regular (circunferencia o elipse), se podría aproximar por otro tipo de modelos basados en formas geométricas no regulares.

- En relación al método de detección, se podría estudiar la construcción de un sistema que analizara la fiabilidad de las estructuras detectadas. Así, por ejemplo, un resultado basado en una tripleta ganadora tendría una alta probabilidad de ser correcta, dado que las tres estructuras así obtenidas cumplen todas las restricciones impuestas por el conocimiento relacional manejado. En cambio, una solución basada en parejas ganadoras, implicará que hay parte de las restricciones que se incumplen y, por tanto, la fiabilidad del resultado ya no es máxima. En última instancia, una solución basada en blobs aislados, implica que no se satisface ninguna de las restricciones y, por tanto, la fiabilidad del resultado es mínima (alguna de las estructuras obtenidas como solución debe ser errónea).
- En cuanto al método de segmentación, se podrían investigar las siguientes líneas: (a) el uso de otras variantes DE o, incluso, otros algoritmos evolutivos en la búsqueda de la solución; (b) la reducción del coste computacional del método mediante la paralelización del AE utilizado; y (c) la incorporación de otras estructuras en el proceso de segmentación simultánea como, por ejemplo, el haz vascular. Esto último debería incrementar la robustez de la búsqueda al estar apoyada en más CRinter-e.
- Finalmente, con carácter general, las dos metodologías propuestas podrían utilizarse para la detección/segmentación de estructuras en otros tipos de imágenes, relacionadas tanto con dominios médicos como no médicos. En el primer caso, por ejemplo, para la detección/segmentación de estructuras y/o órganos anatómicos en radiografía digital o imágenes magnéticas de resonancia y, en el segundo, por ejemplo, para la detección de tipos de componentes electrónicos en imágenes de circuitos impresos. Un requisito importante para aplicar ambas metodologías es que el conocimiento del dominio de aplicación debería contener información topológica relacionada con las estructuras a detectar (CRinter-e).

# Apéndice A

## Publicaciones asociadas a la tesis

Molina, J.M. and Carmona, E.J., “**Localization and Segmentation of the Optic Nerve Head in Eye Fundus Images Using Pyramid Representation and Genetic Algorithms**”, 4th International Work-conference on the Interplay between Natural and Artificial Computation, (IWINAC 2011), May 30-Jun 3, (2011), La Palma, Canary Islands (Spain). **En:** J.M. Ferrández et al (eds.), *Foundations on Natural and Artificial Computation* (Part I), LNCS-6686, Springer-Verlag, pp. 431-440, (2011). **Índices de calidad:** Clase “C” en *Microsoft Academic* (MA); Clase “Work in Progress” en GII-GRIN-SCIE (GGS).



Molina, J.M. and Carmona, E.J., “**Pattern Detection in Images Using LBP-Based Relational Operators**”, 5th International Work-conference on the Interplay between Natural and Artificial Computation, (IWINAC 2013), Jun 10-14, (2013), Mallorca, Balearic Islands (Spain). **En:** J.M. Ferrández et al (eds.), *Natural and Artificial Computation in Engineering and Medical Applications*, LNCS-7931, Springer-Verlag, pp. 11-20, (2013). **Índices de calidad:** Clase “C” en *Microsoft Academic* (MA); Clase “Work in Progress” en GII-GRIN-SCIE (GGS).

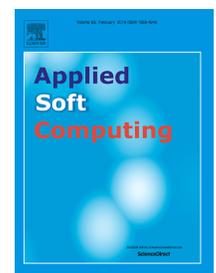


---

Molina-Casado, J.M., Carmona, E.J., and García-Feijoó, J., “**Fast detection of the main anatomical structures in digital retinal images based on intra- and inter-structure relational knowledge**”, *Computer Methods and Programs in Biomedicine*, Vol. 149, pp. 55-68, (2017). **Índices de calidad (2016)**: Factor de impacto (JCR): 2.503 (Q1); Ranking: 21/104 (área “*Computer Science, Theory & Methods*”).



Carmona, E.J. and Molina-Casado, J.M. “**Using Evolutionary Algorithms and Relational Models to Simultaneously Segment the Optic disc and Fovea in Retinal Images**”, *Applied Soft Computing*, (2018). (*En revisión*) **Índices de calidad (2016)**: Factor de impacto (JCR): 3.541 (Q1); Ranking: 21/133 (área “*Computer Science, Artificial Intelligence*”).



# Apéndice B

## Conclusions and future work

In this chapter we present the different conclusions of this thesis, as well as a number of proposals for research lines in the future.

### B.1. Conclusions

In relation to the objectives set out in this thesis, the following conclusions should be highlighted:

- Different models based on intra-structure relational knowledge (intra-SRK models) and inter-structure relational knowledge (inter-SRK models) of the main retinal structures (OD, macula/fovea, vascular bundle) have been designed. The complexity of each of them has depended on the type of task to be solved: detection or segmentation. The objectives 1 and 2 initially stated in this thesis are met.
- Two methodologies have been presented for the detection and segmentation of the main anatomical structures in retinal images. Both methodologies are independent of the original size of the image, they pose the detection/segmentation of structures simultaneously and both of them use two types of relational knowledge: intra and inter-structure. The objectives 3 and 4 initially stated are met.
- A method based on the aforementioned detection methodology has been implemented

and evaluated. Its goal was to simultaneously detect the OD, macula, vascular bundle and network of main retinal vessels. The method is fast, simple and its detection results are competitive in relation to other existing methods in the recent literature, including methods of individual, sequential or simultaneous detection. The main advantage of the simultaneous detection and the joint use of intra and inter-SRK was the reduction in the number of false positives, compared to other variants of this method where the detection was performed individually or sequentially, or where intra-SRK was only used. There is no known method in the related literature that has addressed the challenge of detecting four retinal structures (either sequentially or simultaneously). Maintaining the same configuration of parameters, the method was applied to different public image databases and, in all of them, the detection success rates obtained were consistent, evidencing its robustness. The objective 5 initially stated is met.

- A method based on the aforementioned segmentation methodology has been implemented and evaluated. Its goal was to simultaneously segment the OD and fovea. The segmentation results obtained are competitive in relation to other existing methods in the recent literature, including methods of individual, sequential or simultaneous segmentation. Again, the main advantage of the simultaneous segmentation and the joint use of intra and inter-SRK was the reduction in the number of false positives, compared to other variants of this method where segmentation was performed individually or sequentially, or where only intra-SRK was used. Conversely, the main inconvenient was the computational cost associated with the search for a solution based on the evolution of a population. Maintaining the same configuration of parameters, the method was applied to two public image databases and, in both cases, the detection success rates obtained were consistent, evidencing its robustness. The objective 6 initially stated is met.
- A strategy has been implemented to allow the segmentation method to efficiently use the result obtained by the detection method. The objective 7 initially stated is met.
- A fast and efficient method for the automatic detection of the type of eye (left or right)

shown in a retinal image has been implemented and evaluated. It is based on the use of the intra-SRK vascular bundle model. This method was not initially considered in the objectives of this thesis.

In relation to the hypotheses raised in this thesis, the following conclusions should be noted:

- The results obtained with the implemented methods, both with detection and segmentation, support the following hypothesis: The joint use of intra and inter-SRK in the detection or segmentation process favors the reduction of false positives in the results obtained and, consequently, increases the success rate with respect to those cases in which only intra-SRK is used.
- The results obtained with the implemented methods, both with detection and segmentation, support the following hypothesis: The simultaneous detection/segmentation of  $n$  structures from the joint use of intra and inter-SRK favors the choice of that  $n$ -tuple of structures that best explains both types of knowledge and, therefore, also increases the validity of the detection/segmentation process.
- The results obtained with the segmentation method with detection information support the following hypothesis: the additional incorporation of the detection knowledge (output of the detection process) to the segmentation process favours the success rate. However, the computational cost increases by 28% compared to the original segmentation method.

## B.2. Future work

As future work, it is worth mentioning the possible study of the following lines of research:

- A possible line could be aimed at the creation of more precise models of retinal structures. Thus, for example, given that the contour of the OD does not obey strictly to a regular geometric shape (circumference or ellipse), it could be approximated by other type of models based on non-regular geometric shapes.

- In relation to the detection method, the construction of a system that analyzes the reliability of the detected structures could be studied. Thus, for example, a result based on a winning triplet would have a high probability of being correct, given that the three structures thus obtained fulfill all the restrictions imposed by the managed relational knowledge. In contrast, a solution based on winning couples will imply that part of the restrictions are violated and, therefore, the reliability of the result is no longer maximum. Ultimately, a solution based on isolated blobs implies that none of the constraints is satisfied and, therefore, the reliability of the result is minimal (some of the structures obtained as a solution must be erroneous).
- Regarding the segmentation method, the following lines could be researched: (a) the use of other DE variants or, even, other evolutionary algorithms in the search for the solution; (b) the reduction of the computational cost of the method through the parallelization of the EA used; and (c) the incorporation of other structures in the process of simultaneous segmentation, such as the vascular bundle. The latter should increase the robustness of the search by being supported by more inter-SRK.
- Finally, the two proposed methodologies could be used for the detection/segmentation of structures in other types of images, related to both medical and non-medical domains. In the first case, for example, for the detection/segmentation of structures and/or anatomical organs in digital radiography or magnetic resonance images and, in the second case, for example, for the detection of types of electronic components in printed circuit images. An important requirement to apply both methodologies is that the domain knowledge of the problem should contain topological information related to the structures to be detected (inter-SRK).

# Bibliografía

- [Abdel-Ghafar and Morris, 2007] Abdel-Ghafar, R. A. and Morris, T. (2007). Progress towards automated detection and characterization of the optic disc in glaucoma and diabetic retinopathy. *Medical Informatics and the Internet in Medicine*, 32(1):19–25.
- [Al-Bander et al., 2018] Al-Bander, B., Al-Nuaimy, W., Williams, B. M., and Zheng, Y. (2018). Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc. *Biomedical Signal Processing and Control*, 40:91–101.
- [Allam et al., 2015] Allam, A. M. N., Youssif, A. A., and Ghalwash, A. Z. (2015). Automatic segmentation of optic disc in eye fundus images: A survey. *Electronic Letters on Computer Vision and Image Analysis*, 14(1):1–20.
- [Almazroa et al., 2015] Almazroa, A., Burman, R., Raahemifar, K., and Lakshminarayanan, V. (2015). Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey. *Journal of Ophthalmology*, 180972:1–28.
- [Alonso-Montes et al., 2008] Alonso-Montes, C., Vilarino, D. L., Dudek, P., and Penedo, M. G. (2008). Fast retinal vessel tree extraction: A pixel parallel approach. *International Journal of Circuit Theory and Applications*, 36(5-6):641–651.
- [Aquino, 2014] Aquino, A. (2014). Establishing the macular grading grid by means of fovea centre detection using anatomical-based and visual-based features. *Computers in Biology and Medicine*, 55:61–73.

- [Aquino et al., 2010] Aquino, A., Gegúndez-Arias, M. E., and Marin, D. (2010). Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *IEEE Transactions on Medical Imaging*, 29(11):1860–1869.
- [Aslantas and Tunckanat, 2007] Aslantas, V. and Tunckanat, M. (2007). Differential evolution algorithm for segmentation of wound images. In *IEEE International Symposium on Intelligent Signal Processing*, pages 1–5. IEEE.
- [Atkinson and Mazo, 2011] Atkinson, A. and Mazo, C. (2011). Imaged area of the retina. <https://www.freelists.org/archives/optimal/02-2017/pdf91WmMGLh6Q.pdf>. (último acceso 17.04.18).
- [Ayub et al., 2016] Ayub, J., Ahmad, J., Muhammad, J., Aziz, L., Ayub, S., Akram, U., and Basit, I. (2016). Glaucoma detection through optic disc and cup segmentation using k-mean clustering. In *International Conference on Computing, Electronic and Electrical Engineering*, pages 143–147. IEEE.
- [Besenczi et al., 2016] Besenczi, R., Tóth, J., and Hajdu, A. (2016). A review on automatic analysis techniques for color fundus photographs. *Computational and Structural Biotechnology Journal*, 14:371–384.
- [Beyer and Schwefel, 2002] Beyer, H.-G. and Schwefel, H.-P. (2002). Evolution strategies – a comprehensive introduction. *Natural Computing*, 1(1):3–52.
- [Bharkad, 2017] Bharkad, S. (2017). Automatic segmentation of optic disk in retinal images. *Biomedical Signal Processing and Control*, 31:483–498.
- [Calvo et al., 2011] Calvo, D., Ortega, M., Penedo, M. G., and Rouco, J. (2011). Automatic detection and characterisation of retinal vessel tree bifurcations and crossovers in eye fundus images. *Computer Methods and Programs in Biomedicine*, 103(1):28–38.
- [Cao et al., 2013] Cao, Q., Liu, J., and Zhao, Q. (2013). Fast automatic optic disc localization in retinal images. In *7th International Conference on Image and Graphics*, pages 827–831. IEEE.

- [Carmona et al., 2008] Carmona, E. J., Rincón, M., García-Feijoo, J., and Martínez-de-la Casa, J. M. (2008). Identification of the optic nerve head with genetic algorithms. *Artificial Intelligence in Medicine*, 43:243–259.
- [Chan and Vese, 2001] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.
- [Cheng et al., 2011] Cheng, J., Liu, J., Wong, D. W. K., Yin, F., Cheung, C., Baskaran, M., Aung, T., and Wong, T. Y. (2011). Automatic optic disc segmentation with peripapillary atrophy elimination. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6224–6227. IEEE.
- [Chrástek et al., 2002] Chrástek, R., Wolf, M., Donath, K., Michelson, G., and Niemann, H. (2002). Optic disc segmentation in retinal images. In Meiler, M. et al., editors, *Bildverarbeitung Für Die Medizin*, pages 263–266. Springer, Berlin, Heidelberg.
- [Chutatape, 2006] Chutatape, O. (2006). Fundus foveal localization based on vessel model. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4440–4444. IEEE.
- [Ciulla and Amador, 2003] Ciulla, T. and Amador, A. (2003). Diabetic retinopathy and diabetic macular edema. pathophysiology, screening, and novel therapies. *Diabetes Care*, 26(9):2653–2664.
- [Cox and Wood, 1991] Cox, M. J. and Wood, I. C. (1991). Computer-assisted optic nerve head assessment. *Ophthalmic and Physiological Optic*, 11(1):27–35.
- [Das and Suganthan, 2011] Das, S. and Suganthan, P. N. (2011). Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation*, 15(1):4–31.
- [Dashtbozorg et al., 2015] Dashtbozorg, B., Mendonça, A. M., and Campilho, A. (2015). Optic disc segmentation using the sliding band filter. *Computers in Biology and Medicine*, 56:1–12.

- [Dashtbozorg et al., 2016] Dashtbozorg, B., Zhang, J., Huang, F., and ter Haar Romeny, B. M. (2016). Automatic optic disc and fovea detection in retinal images using super-elliptical convergence index filters. In Campilho, A. and Karray, F., editors, *Image Analysis and Recognition*, volume 9730 of *LNCS*, pages 697–706. Springer, Cham.
- [Díaz-Pernil et al., 2016] Díaz-Pernil, D., Fondón, I., Peña-Cantillana, F., and Gutiérrez-Naranjo, M. A. (2016). Fully automatized parallel segmentation of the optic disc in retinal fundus images. *Pattern Recognition Letters*, 83:99–107.
- [de la Fuente-Arriaga et al., 2014] de la Fuente-Arriaga, J. A., Felipe-Riverón, E. M., and Garduño Calderón, E. (2014). Application of vascular bundle displacement in the optic disc for glaucoma detection using fundus images. *Computers in Biology and Medicine*, 47:27–35.
- [Decenciere et al., 2014] Decenciere, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Tro-  
ne, C., Gain, P., Ordonez, R., Massin, P., Erginay, A., Charton, B., and Klein, J. (2014). Feedback on a publicly distributed database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234.
- [DIARETDB1, 2015] DIARETDB1 (2015). The diabetic retinopathy database. <http://www.it.lut.fi/project/imageret/diaretdb1/>. (último acceso 17.04.18).
- [DRIONS-DB, 2013] DRIONS-DB (2013). Digital retinal images for optic nerve segmentation database. <http://www.ia.uned.es/personal/ejcarmona/DRIONS-DB.html>. (último acceso 17.04.18).
- [DRIVE, 2004] DRIVE (2004). Drive: Digital retinal images for vessel extraction. <http://www.isi.uu.nl/Research/Databases/DRIVE/>. (último acceso 17.04.18).
- [Fischler and Bolles, 1981] Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395.

- [Fleming et al., 2007] Fleming, A., Goatman, K., Philip, S., Olson, J., and Sharp, P. (2007). Automatic detection of retinal anatomy to assist diabetic retinopathy screening. *Physics in Medicine and Biology*, 52(2):331–345.
- [Foracchia et al., 2004] Foracchia, M., Grisan, E., and Ruggeri, A. (2004). Detection of optic disc in retinal images by means of a geometrical model of vessel structure. *IEEE Transactions on Medical Imaging*, 23(10):1189–1195.
- [Gagnon et al., 2001] Gagnon, L., Lalonde, M., and Beaulieu, M. (2001). Procedure to detect anatomical structures in optical fundus images. In *SPIE Medical Imaging*, volume 4322, pages 1218–1225.
- [Gegundez et al., 2013] Gegundez, M. E., Marin, D., Bravo, J. M., and Suero, A. (2013). Locating the fovea center position in digital fundus images using thresholding and feature extraction techniques. *Computerized Medical Imaging and Graphics*, 37(5-6):386–393.
- [Giachetti et al., 2014] Giachetti, A., Ballerini, L., and Trucco, E. (2014). Accurate and reliable segmentation of the optic disc in digital fundus images. *Journal of Medical Imaging*, 1(2):024001.
- [Giachetti et al., 2013] Giachetti, A., Ballerini, L., Trucco, E., and Wilson, P. (2013). The use of radial symmetry to localize retinal landmarks. *Computerized Medical Imaging and Graphics*, 37(5-6):369–376.
- [Giachetti et al., 2011] Giachetti, A., Chin, K. S., Trucco, E., Cobb, C., and Wilson, P. J. (2011). Multiresolution localization and segmentation of the optical disc in fundus images using inpainted background and vessel information. In *18th IEEE International Conference on Image Processing*, pages 2145–2148. IEEE.
- [Girard et al., 2016] Girard, F., Kavalec, C., Grenier, S., Ben Tahar, H., and Chériet, F. (2016). Simultaneous macula detection and optic disc boundary segmentation in retinal fundus images. In *SPIE Medical Imaging*, volume 9784, pages 97841F1–97841F9.

- [Haleem et al., 2013] Haleem, M. S., Han, L., van Hemert, J., and Li, B. (2013). Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: A review. *Computerized Medical Imaging and Graphics*, 37(7-8):581–596.
- [Hansen, 2006] Hansen, N. (2006). *The CMA Evolution Strategy: A Comparing Review*, pages 75–102. Springer Berlin Heidelberg.
- [Holland, 1975] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. The University of Michigan Press.
- [Hooshyar and Khayati, 2010] Hooshyar, S. and Khayati, R. (2010). Retina vessel detection using fuzzy ant colony algorithm. In *Canadian Conference on Computer and Robot Vision*, pages 239–244. IEEE.
- [Hoover and Goldbaum, 2003] Hoover, A. and Goldbaum, M. (2003). Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging*, 22(8):951–958.
- [Hoover et al., 2000] Hoover, A., Kouznetsova, V., and Goldbaum, M. (2000). Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210.
- [Hubbard et al., 1999] Hubbard, L. D., Brothers, R. J., King, W. N., Clegg, L. X., Klein, R., Cooper, L. S., Sharrett, A. R., Davis, M. D., and Cai, J. (1999). Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study. *Ophthalmology*, 106(12):2269–2280.
- [Jelinek et al., 2005] Jelinek, H. F., Depardieu, C., Lucas, C., Cornforth, D. J., Huang, W., and Cree, M. J. (2005). Towards vessel characterisation in the vicinity of the optic disc in digital retinal images. In McCane, B., editor, *Conference of Image and Vision Computing*, pages 2–7.

- 
- [Jonas et al., 2015] Jonas, R. A., Wang, Y. X., Yang, H., Li, J. J., Xu, L., Panda-Jonas, S., and Jonas, J. B. (2015). Optic disc - fovea angle: The beijing eye study 2011. *PLOS ONE*, 10(11):1–10.
- [Joshi et al., 2010a] Joshi, G. D., Gautam, R., Sivaswamy, J., and Krishnadas, S. R. (2010a). Robust optic disk segmentation from colour retinal images. In *7th Indian Conference on Computer Vision, Graphics and Image Processing*, pages 330–336. ACM.
- [Joshi et al., 2010b] Joshi, G. D., Sivaswamy, J., Karan, K., and Krishnadas, S. R. (2010b). Optic disk and cup boundary detection using regional information. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 948–951. IEEE.
- [Kande et al., 2008] Kande, G. B., Subbaiah, P. V., and Savithri, T. S. (2008). Segmentation of exudates and optic disk in retinal images. In *6th Indian Conference on Computer Vision, Graphics Image Processing*, pages 535–542. IEEE.
- [Kao et al., 2014] Kao, E. F., Lin, P. C., Chou, M. C., Jaw, T. S., and Liu, G. C. (2014). Automated detection of fovea in fundus images based on vessel-free zone and adaptive gaussian template. *Computer Methods and Programs in Biomedicine*, 117(2):92–103.
- [Kauppi et al., 2007] Kauppi, T., Kalesnykiene, V., Kamarainen, J., Lensu, L., Sorri, I., Raniinen, A., Voutilainen, R., Uusitalo, H., Kälviäinen, H., and Pietilä, J. (2007). The diaretdb1 diabetic retinopathy database and evaluation protocol. In *11th Conference on Medical Image Understanding and Analysis*, pages 61–65. BMVC.
- [Kim et al., 2007] Kim, S. K., Kong, H. J., Seo, J. M., Cho, B. J., Park, K. H., Hwang, J. M., Kim, D. M., Chung, H., and Kim, H. C. (2007). Segmentation of optic nerve head using warping and ransac. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 900–903. IEEE.
- [Kusumandari et al., 2015] Kusumandari, D. E., ArisMunandar, and Redhyka, G. G. (2015). The comparison of gvf snake active contour method and ellipse fit in optic disc detection for glaucoma diagnosis. In *International Conference on Automation, Cognitive Science,*

*Optics, Micro Electro-Mechanical System, and Information Technology*, pages 123–126. IEEE.

[Lalonde et al., 2001] Lalonde, M., Beaulieu, M., and Gagnon, L. (2001). Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching. *IEEE Transaction on Medical Imaging*, 20(11):1193–1200.

[Lee et al., 2010] Lee, S., Abramoff, M., and Reinhardt, J. (2010). Retinal atlas statistics from color fundus images. In *SPIE Medical Imaging*, volume 7623, pages 762310–762319.

[Lee and Brady, 1991] Lee, S. and Brady, M. (1991). Optic disk boundary detection. In *British Machine Vision Conference*, pages 359–362. Springer, London.

[Li and Chutatape, 2004] Li, H. and Chutatape, O. (2004). Automated feature extraction in color retinal images by a model based approach. *IEEE Transactions on Biomedical Engineering*, 51(2):246–254.

[Lowell et al., 2004] Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., and Fletcher, E. (2004). Optic nerve head segmentation. *IEEE Transactions on Medical Imaging*, 23(2):256–264.

[Lu, 2011] Lu, S. (2011). Accurate and efficient optic disc detection and segmentation by a circular transformation. *IEEE Transactions on Medical Imaging*, 30(12):2126–2133.

[Lu and Lim, 2010a] Lu, S. and Lim, J. H. (2010a). Automatic macula detection from retinal images by a line operator. In *17th IEEE International Conference on Image Processing*, pages 4073–4076. IEEE.

[Lu and Lim, 2010b] Lu, S. and Lim, J. H. (2010b). Automatic optic disc detection through background estimation. In *17th IEEE International Conference on Image Processing*, pages 833–836. IEEE.

- [Mahesh, 2013] Mahesh, K. (2013). A survey of automated techniques for retinal disease identification in diabetic retinopathy. *International Journal of Advancements in Research & Technology*, 2(5):199–216.
- [Maheshwari et al., 2017] Maheshwari, S., Pachori, R. B., and Acharya, U. R. (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):803–813.
- [Mahfouz and Fahmy, 2010] Mahfouz, A. E. and Fahmy, A. S. (2010). Fast localization of the optic disc using projection of image features. *IEEE Transactions on Image Processing*, 19(12):3285–3289.
- [Marin et al., 2011] Marin, D., Aquino, A., Gegundez-Arias, M. E., and Bravo, J. M. (2011). A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Transactions on Medical Imaging*, 30(1):146–158.
- [Marin et al., 2015] Marin, D., Gegundez-Arias, M., Suero, A., and Bravo, J. (2015). Obtaining optic disc center and pixel region by automatic thresholding methods on morphologically processed fundus images. *Computer Methods and Programs in Biomedicine*, 118(2):173–185.
- [Mary et al., 2016] Mary, M. C. V. S., Rajsingh, E. B., and Naik, G. R. (2016). Retinal fundus image analysis for diagnosis of glaucoma: A comprehensive survey. *IEEE Access*, 4:4327–4354.
- [McInerney and Terzopoulos, 1995] McInerney, T. and Terzopoulos, D. (1995). Topologically adaptable snakes. In *IEEE International Conference on Computer Vision*, pages 840–845. IEEE.

- [Medhi and Dandapat, 2016] Medhi, J. and Dandapat, S. (2016). An effective fovea detection and automatic assessment of diabetic maculopathy in color fundus images. *Computers in Biology and Medicine*, 74:30–44.
- [Mendels et al., 1999] Mendels, F., Heneghan, C., and Thiran, J. P. (1999). Identification of the optic disk boundary in retinal images using active contours. In *Irish Machine Vision and Image Processing Conference*, volume 5112, pages 103–115. IEEE.
- [Mesejo et al., 2016] Mesejo, P., Ibañez, Ó., Cordón, Ó., and Cagnoni, S. (2016). A survey on image segmentation using metaheuristic-based deformable models: state of the art and critical analysis. *Applied Soft Computing*, 44:1–29.
- [Mesejo et al., 2013] Mesejo, P., Ugolotti, R., Cunto, F. D., Giacobini, M., and Cagnoni, S. (2013). Automatic hippocampus localization in histological images using differential evolution-based deformable models. *Pattern Recognition Letters*, 34(3):299–307.
- [MESSIDOR, 2012] MESSIDOR (2012). Messidor database. <http://messidor.crihan.fr>. (último acceso 17.04.18).
- [Michalewicz, 1996] Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs (3rd Ed.)*. Springer-Verlag.
- [Mittal and Sivaswamy, 2015] Mittal, G. and Sivaswamy, J. (2015). Optic disk and macula detection from retinal images using generalized motion pattern. In *5th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, pages 1–4. IEEE.
- [Molina and Carmona, 2011] Molina, J. M. and Carmona, E. (2011). Localization and segmentation of the optic nerve head in eye fundus images using pyramid representation and genetic algorithms. In Ferrández, J. M. et al., editors, *Foundations on Natural and Artificial Computation*, volume 6686 of *LNCS*, pages 431–440. Springer, Berlin, Heidelberg.

- [Morales et al., 2013] Morales, S., Naranjo, V., Angulo, J., and Alcañiz, M. L. (2013). Automatic detection of optic disc based on pca and mathematical morphology. *IEEE Transactions on Medical Imaging*, 32(4):786–796.
- [Morris and Donnison, 1999] Morris, T. D. and Donnison, C. (1999). Identifying the neuroretinal rim boundary using dynamic contours. *Image and Vision Computing*, 17(3):169–174.
- [Nekovei and Sun, 1995] Nekovei, R. and Sun, Y. (1995). Back-propagation network and its configuration for blood vessel detection in angiograms. *IEEE Transactions on Neural Networks*, 6(1):64–72.
- [Niemeijer et al., 2007] Niemeijer, M., Abramoff, M., and van Ginneken, B. (2007). Segmentation of the optic disc, macula and vascular arch in fundus photographs. *IEEE Transactions on Medical Imaging*, 26(1):116–127.
- [Niemeijer et al., 2009] Niemeijer, M., Abramoff, M. D., and van Ginneken, B. (2009). Fast detection of the optic disc and fovea in color fundus photographs. *Medical Image Analysis*, 13(6):859–870.
- [Noor et al., 2013] Noor, N. M., Khalid, N. E. A., and Ariff, N. M. (2013). Optic cup and disc color channel multi-thresholding segmentation. In *IEEE International Conference on Control System, Computing and Engineering*, pages 530–534. IEEE.
- [Novo et al., 2009] Novo, J., Penedo, M. G., and Santos, J. (2009). Localisation of the optic disc by means of ga-optimised topological active nets. *Image and Vision Computing*, 27(10):1572–1584.
- [Nyul, 2009] Nyul, L. G. (2009). Retinal image analysis for automated glaucoma risk evaluation. In *SPIE Medical Imaging*, volume 7497, pages 74971C1–74971C9.
- [ONHSD, 2013] ONHSD (2013). Optic nerve head segmentation dataset. <http://www.aldiri.info/Image%20Datasets/ONHSD.aspx>. (último acceso 17.04.18).

- [Osareh, 2004] Osareh, A. (2004). *Automated Identification of Diabetic Retinal Exudates and the Optic Disc*. PhD thesis, University of Bristol.
- [Osareh et al., 2002a] Osareh, A., Mirmehdi, M., Thomas, B., and Markham, R. (2002a). Colour morphology and snakes for optic disc localisation. In *6th Medical Image Understanding and Analysis Conference*, pages 21–24. BMVA Press.
- [Osareh et al., 2002b] Osareh, A., Mirmehdi, M., Thomas, B., and Markham, R. (2002b). Comparison of colour spaces for optic disc localisation in retinal images. In *Object recognition supported by user interaction for service robots*, volume 1, pages 743–746.
- [Pallawala et al., 2004] Pallawala, P. M. D. S., Hsu, W., Lee, M. L., and Eong, K.-G. A. (2004). Automated optic disc localization and contour detection using ellipse fitting and wavelet transform. In Pajdla, T. and Matas, J., editors, *8th European Conference on Computer Vision*, pages 139–151. Springer, Berlin, Heidelberg.
- [Petsatodis et al., 2004] Petsatodis, T. S., Diamantis, A., and Syrcos, G. (2004). A complete algorithm for automatic human recognition based on retina vascular network characteristics. In *1st International Scientific Conference e RA*, pages 41–46.
- [Price et al., 2005] Price, K., Storn, R. M., and Lampinen, J. A. (2005). *Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series)*. Springer-Verlag New York, Inc.
- [Qureshi et al., 2012] Qureshi, R. J., Kovacs, L., Harangi, B., Nagy, B., Peto, T., and Hajdu, A. (2012). Combining algorithms for automatic detection of optic disc and macula in fundus images. *Computer Vision and Image Understanding*, 116(1):138–145.
- [Raja and Ravichandran, 2014] Raja, J. B. and Ravichandran, C. G. (2014). Automatic localization of fovea in retinal images based on mathematical morphology and anatomic structures. *International Journal of Engineering & Technology*, 6(5):2171–2183.

- [Ricci and Perfetti, 2007] Ricci, E. and Perfetti, R. (2007). Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging*, 26(10):1357–1365.
- [Rohrschneider, 2004] Rohrschneider, K. (2004). Determination of the location of the fovea on the fundus. *Investigative Ophthalmology & Visual Science*, 45(9):3257–3258.
- [Roy et al., 2016] Roy, P. K., Chakravorty, R., Sedai, S., Mahapatra, D., and Garnavi, R. (2016). Automatic eye type detection in retinal fundus image using fusion of transfer learning and anatomical features. In *International Conference on Digital Image Computing: Techniques and Applications*, pages 1–7. IEEE.
- [Roychowdhury et al., 2016] Roychowdhury, S., Koozekanani, D. D., Kuchinka, S. N., and Parhi, K. K. (2016). Optic disc boundary and vessel origin segmentation of fundus images. *IEEE Journal of Biomedical and Health Informatics*, 20(6):1562–1574.
- [Ruggeri et al., 2007] Ruggeri, A., Grisan, E., and De Luca, M. (2007). An automatic system for the estimation of generalized arteriolar narrowing in retinal images. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6463–6466. IEEE.
- [Salazar-Gonzalez et al., 2014] Salazar-Gonzalez, A., Kaba, D., Li, Y., and Liu, X. (2014). Segmentation of the blood vessels and optic disk in retinal images. *IEEE Journal of Biomedical and Health Informatics*, 18(6):1874–1886.
- [Saraswat et al., 2013] Saraswat, M., Arya, K. V., and Sharma, H. (2013). Leukocyte segmentation in tissue images using differential evolution algorithm. *Swarm and Evolutionary Computation*, 11:46–54.
- [Savitzky and Golay, 1964] Savitzky, A. and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639.
- [Schwiegerling, 2004] Schwiegerling, J. (2004). *Field guide to visual and ophthalmic optics*. Bellingham, WA:SPIE Press.

- [Sekhar et al., 2008] Sekhar, S., Al-Nuaimy, W., and Nandi, A. K. (2008). Automated localization of optic disc and fovea in retinal fundus images. In *16th European Signal Processing Conference*, pages 1–5. IEEE.
- [Siddalingaswamy and Prabhu, 2010] Siddalingaswamy, P. C. and Prabhu, K. G. (2010). Automatic localization and boundary detection of optic disc using implicit active contours. *International Journal of Computer Applications*, 1(6):1–5.
- [Sinthanayothin et al., 1999] Sinthanayothin, C., Boyce, J., Cook, H., and Williamson, T. (1999). Automated localisation of the optic disc, fovea, retinal blood vessels and pathologies from digital colour fundus images. *The British Journal of Ophthalmology*, 83(8):902–910.
- [Soares et al., 2006] Soares, J. V. B., Leandro, J. J. G., Cesar, R. M., Jelinek, H. F., and Cree, M. J. (2006). Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 25(9):1214–1222.
- [Staal et al., 2004] Staal, J. J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and van Ginneken, B. (2004). Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509.
- [Storn and Price, 1997] Storn, R. and Price, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359.
- [Tan et al., 2017] Tan, J. H., Acharya, U. R., Bhandary, S. V., Chua, K. C., and Sivaprasad, S. (2017). Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *Journal of Computational Science*, 20:70–79.
- [Tan et al., 2009] Tan, N. M., Liu, J., Wong, D. W. K., Lim, J. H., Li, H., Patil, S. B., Yu, W., and Wong, T. Y. (2009). Automatic detection of left and right eye in retinal fundus images. In Lim, C. T. and Goh, J. C. H., editors, *13th International Conference on Biomedical Engineering*, volume 23, pages 610–614. Springer, Berlin, Heidelberg.

- [Tang et al., 2006] Tang, Y., Li, X., von Freyberg, A., and Goch, G. (2006). Automatic segmentation of the papilla in a fundus image based on the c-v model and a shape restraint. In *Proceedings of the 18th International Conference on Pattern Recognition*, volume 1, pages 183–186. IEEE.
- [Tang et al., 2003] Tang, Y., von Freyberg, A., Selzer, H., and Goch, G. (2003). Automatic papilla shape detection in stereo retinal images based on image fusion for computer aided glaucoma diagnostics. In *17th IMEKO World Congress, Metrology in the 3rd Millenium*.
- [Taylor and Keeffe, 2001] Taylor, H. R. and Keeffe, J. E. (2001). World blindness: a 21st century perspective. *British Journal of Ophthalmology*, 85(3):261–266.
- [Tobin et al., 2007] Tobin, K., Chaum, E., Govindasamy, V., and Karnowski, T. (2007). Detection of anatomic structures in human retinal imagery. *IEEE Transactions on Medical Imaging*, 26(12):1729–1739.
- [Ugolotti et al., 2013] Ugolotti, R., Nashed, Y. S., Mesejo, P., Ivekovic, S., Mussi, L., and Cagnoni, S. (2013). Particle swarm optimization and differential evolution for model-based object detection. *Applied Soft Computing*, 13(6):3092–3105.
- [Vallabha et al., 2004] Vallabha, D., Dorairaj, R., Namuduri, K., and Thompson, H. (2004). Automated detection and classification of vascular abnormalities in diabetic retinopathy. In *13th IEEE Signals, Systems and Computers, 2004*, volume 2, pages 1625–1629. IEEE.
- [Walter and Klein, 2005] Walter, T. and Klein, J.-C. (2005). Automatic analysis of color fundus photographs and its application to the diagnosis of diabetic retinopathy. In Suri, J. S. et al., editors, *Handbook of Biomedical Image Analysis*, ITBE, pages 315–368. Springer, Boston, MA.
- [Welfer et al., 2010] Welfer, D., Scharcanski, J., Kitamura, C. M., Dal Pizzol, M. M., Ludwig, L. W., and Marinho, D. R. (2010). Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach. *Computers in Biology and Medicine*, 40:124–137.

- [Welfer et al., 2011] Welfer, D., Scharcanski, J., and Marinho, D. (2011). Fovea center detection based on the retina anatomy and mathematical morphology. *Computer Methods and Programs in Biomedicine*, 104(3):397–409.
- [Williams and Wilkinson, 1992] Williams, T. and Wilkinson, J. (1992). Position of the fovea centralis with respect to the optic nerve head. *Optometry & Vision Science*, 69(5):369–377.
- [Wong et al., 2004] Wong, T. Y., Klein, R., Sharrett, A. R., Duncan, B. B., Couper, D. J., Klein, B. E. K., Hubbard, L. D., Nieto, F. J., et al. (2004). Retinal arteriolar diameter and risk for hypertension. *Annals of Internal Medicine*, 140(4):248–255.
- [Xiong and Li, 2016] Xiong, L. and Li, H. (2016). An approach to locate optic disc in retinal images with pathological changes. *Computerized Medical Imaging and Graphics*, 47:40–50.
- [Xu, 2010] Xu, X. (2010). Simultaneous automatic detection of optic disc and fovea. Master’s thesis, University of Iowa.
- [Ying et al., 2007] Ying, H., Zhang, M., and Liu, J.-C. (2007). Fractal-based automatic localization and segmentation of optic disc in retinal images. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4139–4141. IEEE.
- [Youssif et al., 2008] Youssif, A. A. H. A. R., Ghalwash, A. Z., and Ghoneim, A. A. S. A. R. (2008). Optic disc detection from normalized digital fundus images by means of a vessels’ direction matched filter. *IEEE Transactions on Medical Imaging*, 27(1):11–18.
- [Yu et al., 2011] Yu, H., Barriga, E., Agurto, C., Echegaray, S., Pattichis, M., Zamora, G., Bauman, W., and Soliz, P. (2011). Fast localization of optic disc and fovea in retinal images for eye disease screening. In *SPIE Medical Imaging*, volume 7963, pages 796317–796329.
- [Yu et al., 2012] Yu, H., Barriga, E. S., Agurto, C., Echegaray, S., Pattichis, M. S., Bauman, W., and Soliz, P. (2012). Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets. *IEEE Transactions on Information Technology in Biomedicine*, 16(4):644–657.

[Zhu et al., 2010] Zhu, X., Rangayyan, R. M., and Ells, A. L. (2010). Detection of the optic nerve head in fundus images of the retina using the hough transform for circles. *Journal of Digital Imaging*, 23(3):332–341.