

Trabajo de Fin de Máster

**Segmentación de imágenes utilizando  
el operador *Linear Binary Pattern*  
basado en pirámide. Aplicación a la  
segmentación de vasos sanguíneos en  
imágenes de retina.**

Fernando Castellanos Sánchez

Tutor: Enrique J. Carmona Suárez

Septiembre 2022



Universidad Nacional de Educación a Distancia  
Escuela Técnica Superior de Ingeniería Informática  
Máster Universitario en Investigación en Inteligencia Artificial



# Resumen

En este *Trabajo de Fin de Máster* se lleva a cabo el análisis de una metodología de segmentación de imágenes que no se encuentre basada en técnicas de *Deep Learning*, de forma que, sin emplear recursos costosos como una *GPU*, y sin necesidad de una base de datos de gran tamaño, sea capaz de proporcionar resultados competitivos para aquellos problemas en los que la máxima precisión no es requerida.

Con dicha finalidad, se propone la generación de características discriminantes a través del operador *LBP* basado en pirámide de imagen como medio para describir las vecindades de los píxeles en regiones de múltiples dimensiones. Entre dichas características se encuentran las proporcionadas por diferentes variantes del operador *LBP* aplicadas a varias escalas de manera simultánea. Descritas las características de cada uno de los píxeles, se lleva a cabo el entrenamiento de un clasificador «tradicional» capaz de clasificar los píxeles de nuevas imágenes.

Como aplicación, se emplean bases de imágenes de fondo de ojo con el fin de segmentar los vasos sanguíneos. Para ello, se utilizan tres bases de datos diferentes sobre las que existe literatura previa que contrasta su utilización y establece una referencia sobre la que comparar los resultados. De esta forma, se alcanzan finalmente resultados que pueden ser considerados competitivos con la literatura más reciente.

**Palabras Clave:** Local Binary Pattern, Pirámide de imagen, Segmentación de imágenes, Clasificación, Computación evolutiva, Light Gradient Boosting Machine, Imágenes de retina



# Abstract

In this Master Thesis we analyze an image segmentation methodology that is able to provide competitive results for those problems where maximum accuracy is not required and Deep Learning techniques cannot be applied due to the high cost of resources such as GPUs or the need for a large database.

To that end, the generation of discriminant features through the image pyramid-based LBP operator is proposed as a means of describing pixel neighborhoods in multi-dimensional regions. Among such features are those provided by different versions of the LBP operator applied at several scales simultaneously. Once the individual pixel features are described, a "traditional" classifier capable of classifying pixels of new images is trained.

As an application, retinal image databases are used to segment blood vessels. For this purpose, three different databases are analyzed. Their use is contrasted in the literature. After comparing the results, a high accuracy is achieved without such a high computational cost as in the use of Deep Learning.

**Keywords:** Local Binary Pattern, Pyramid, Image segmentation, Classification, Evolutionary Computing, Light Gradient Boosting Machine, Retinal Images



# Índice general

<b>1. Introducción</b>	<b>9</b>
1.1. Motivación . . . . .	10
1.2. Definición del problema . . . . .	11
1.3. Limitaciones . . . . .	12
1.4. Hipótesis y objetivos . . . . .	12
1.5. Estructura del documento . . . . .	13
<b>2. Estado de la Cuestión</b>	<b>15</b>
2.1. Trabajos relacionados con el operador <i>LBP</i> y la pirámide de imagen .	15
2.2. Análisis de trabajos relacionados con la segmentación de vasos en retinografías . . . . .	20
2.2.1. Técnicas de aprendizaje no supervisado . . . . .	20
2.2.2. Técnicas de aprendizaje supervisado . . . . .	21
<b>3. Conocimiento previo</b>	<b>25</b>
3.1. <i>Linear Binary Pattern (LBP)</i> . . . . .	25
3.1.1. Limitaciones . . . . .	26
3.1.2. Variantes del operador <i>LBP</i> . . . . .	27
3.2. Pirámide de imagen . . . . .	29
3.3. Clasificador <i>LightGBM</i> . . . . .	29
3.4. Computación evolutiva . . . . .	32
<b>4. Metodología</b>	<b>33</b>
4.1. Bases de datos . . . . .	33
4.2. Algoritmo de segmentación de vasos sanguíneos . . . . .	35
4.2.1. Definición de la estructura del dataset . . . . .	35
4.2.2. Esquema general de funcionamiento del algoritmo . . . . .	39
4.2.3. Otras etapas adicionales . . . . .	44
<b>5. Experimentos, resultados y discusión</b>	<b>53</b>
5.1. Consideraciones sobre los conjuntos de entrenamiento, de validación y de test . . . . .	54
5.2. Métricas de evaluación de los resultados . . . . .	55
5.3. Experimentos realizados . . . . .	56
5.3.1. Pirámide de imagen con el operador <i>LBP</i> . . . . .	57
5.3.2. Pirámide de imagen con diferentes variantes del operador <i>LBP</i>	58
5.3.3. Introducción de la intensidad del nivel de gris como característica	60

5.3.4.	Utilización de la escala 2:1 . . . . .	61
5.3.5.	Selección final de características basadas en el operador <i>LBP</i> .	62
5.3.6.	Utilización del canal verde . . . . .	66
5.3.7.	Sintonización de la etapa de preprocesamiento . . . . .	66
5.3.8.	Generación de características mediante sintonización de nú- cleos de convolución . . . . .	68
5.3.9.	Sintonización de parámetros mediante <i>Grid Search</i> . . . . .	72
5.3.10.	Etapas de postprocesamiento . . . . .	72
5.4.	Resultados de evaluación sobre imágenes de test . . . . .	74
5.4.1.	Base de datos <i>DRIVE</i> . . . . .	77
5.4.2.	Base de datos <i>CHASE</i> . . . . .	82
5.4.3.	Base de datos <i>STARE</i> . . . . .	85
<b>6.</b>	<b>Conclusiones y trabajo futuro</b>	<b>91</b>
6.1.	Conclusiones . . . . .	91
6.2.	Trabajo futuro . . . . .	93
<b>A.</b>	<b>Apéndice</b>	<b>95</b>
A.1.	Resultados gráficos del proceso de constitución de la metodología . .	95
	<b>Referencias</b>	<b>101</b>

# Índice de figuras

3.1.	Ejemplo de cálculo del operador <i>LBP</i> para un píxel de análisis (píxel central), considerando el caso más simple ( $R=1$ y $P=8$ ). . . . .	26
3.2.	Continuación del ejemplo de la Fig. 3.1 para obtener la versión invariante a la rotación para el caso $R = 1$ y $P = 8$ . Las ocho codificaciones binarias mostradas serían equivalentes y 45 el valor finalmente escogido. . . . .	27
3.3.	Representación [1] de la totalidad de posibles patrones del operador <i>RI</i> con $R = 1$ y $P = 8$ . En la primera fila se muestran los patrones uniformes de la variante <i>RIU2</i> . . . . .	28
3.4.	Diagramas de construcción de cada uno de los árboles de decisión en los algoritmos <i>XGBoost</i> y <i>LightGBM</i> , imágenes superior e inferior respectivamente. Fuente: [2]. . . . .	32
4.1.	Ejemplo de imagen del dataset <i>DRIVE</i> junto con su respectivo <i>ground truth</i> . . . . .	34
4.2.	Ejemplo de imagen del dataset <i>STARE</i> junto con su respectivo <i>ground truth</i> . . . . .	34
4.3.	Ejemplo de imagen del dataset <i>CHASE</i> junto con su respectivo <i>ground truth</i> . . . . .	35
4.4.	Diagrama del algoritmo de generación del conjunto de datos. En verde se encuentran las etapas de la Sec. 4.2.3. . . . .	36
4.5.	Diagrama de entrenamiento y evaluación del clasificador. En verde se encuentran las etapas de la Sec. 4.2.3. . . . .	37
4.6.	Generación del conjunto de datos estructurado mediante imágenes a dos escalas diferentes ( $1 : 1$ y $1 : 2$ ). . . . .	38
4.7.	Generación del conjunto de datos para una escala $2 : 1$ . . . . .	39
4.8.	Ejemplo de resultado de la primera operación de filtrado para la reducción de ruido en el proceso de preprocesamiento. . . . .	40
4.9.	Ejemplo de resultado tras la normalización de la imagen en el proceso de preprocesamiento. . . . .	41
4.10.	Ejemplo de resultado de la segunda operación de filtrado para la reducción de ruido en el proceso de preprocesamiento. . . . .	42
4.11.	Ejemplo de predicción de píxeles etiquetados como vasos durante la fase de predicción. . . . .	43
4.12.	Resultado del preprocesamiento sobre el canal verde en una de las imágenes del dataset <i>DRIVE</i> . . . . .	47

4.13. Probabilidad asociada a cada píxel de formar parte de un vaso sanguíneo. Los tonos más blancos indican una mayor probabilidad. . . . .	51
5.1. Representación gráfica de la predicción sobre una imagen de validación empleando las características generadas por 4 escalas. Los píxeles rojos han sido etiquetados como vasos sanguíneos. . . . .	58
5.2. Resultado gráfico de la aplicación de las diferentes variantes del operador <i>LBP</i> a escala 1 : 1. . . . .	59
5.3. Representación gráfica de la predicción sobre una imagen de validación empleando las características generadas por 4 escalas con múltiples variantes del operador <i>LBP</i> simultáneamente. . . . .	60
5.4. Representación gráfica de la predicción sobre una imagen de validación con la adición de la intensidad. . . . .	61
5.5. Resultado gráfico de la aplicación de la característica adicional correspondiente a la escala 2 : 1. . . . .	63
5.6. Resultado gráfico de la predicción con las características seleccionadas mediante validación cruzada. . . . .	65
5.7. Resultado gráfico de la predicción sobre la primera imagen de validación empleando el canal verde. . . . .	67
5.8. Curva de progreso de la evaluación de la población durante el ajuste del preprocesamiento en la base de datos <i>DRIVE</i> en la ejecución del <i>algoritmo evolutivo</i> . . . . .	68
5.9. Resultado gráfico de la predicción empleando el preprocesamiento con parámetros sintonizados mediante un <i>algoritmo evolutivo</i> . . . . .	69
5.10. Curva de progreso de la evaluación de la población durante el ajuste de las convoluciones en la base de datos <i>DRIVE</i> en la ejecución del <i>algoritmo evolutivo</i> . . . . .	70
5.11. Resultado gráfico de la predicción sobre la primera imagen de validación de <i>DRIVE</i> empleando adicionalmente convoluciones. . . . .	70
5.12. Representación de las convoluciones sobre una de las imágenes de validación del dataset <i>DRIVE</i> . . . . .	71
5.13. Resultado gráfico de la predicción sobre la primera imagen de validación de <i>DRIVE</i> en la utilización de parámetros optimizados. . . . .	73
5.14. Resultado gráfico de la predicción sobre la primera imagen de validación <i>DRIVE</i> tras la aplicación del postprocesamiento. . . . .	74
5.15. Comparación de continuidad en un vaso sanguíneo. La imagen de la derecha incluye el postprocesamiento. . . . .	74
5.16. Comparación relativa a la reducción de ruido. La imagen de la derecha incluye el postprocesamiento. . . . .	75
5.17. Comparación de las intersecciones entre vasos. La imagen de la derecha incluye el postprocesamiento. . . . .	75
5.18. Comparación del nervio óptico. La imagen de la derecha incluye el postprocesamiento. . . . .	76

5.19. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>DRIVE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	78
5.20. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>DRIVE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	79
5.21. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>DRIVE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	80
5.22. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>DRIVE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	81
5.23. Regiones con patologías de las imágenes de <i>DRIVE</i> . Se tiene de color rojo aquellos píxeles etiquetados como vasos sanguíneos en el proceso de predicción. . . . .	82
5.24. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>CHASE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	84
5.25. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>CHASE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	85
5.26. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>STARE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	87
5.27. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>STARE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	88
5.28. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>STARE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	89
5.29. Representación de las predicciones de los vasos sanguíneos en el conjunto de test de <i>STARE</i> sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo. . .	90
5.30. Regiones de las imágenes <i>im0044</i> y <i>im0139</i> (izquierda y derecha respectivamente) sobre las que se observa un empeoramiento del poder predictivo debido a las patologías presentes. . . . .	90
A.1. Imagen « <i>35_training.tif</i> » del dataset <i>DRIVE</i> de manera previa a la aplicación del preprocesamiento. . . . .	95
A.2. Imagen « <i>35_training.tif</i> » del dataset <i>DRIVE</i> tras la aplicación del preprocesamiento manual. . . . .	96
A.3. Representación de la predicción de los vasos sanguíneos (píxeles rojos) al aplicar el operador <i>LBP</i> a diferentes escalas en una imagen de validación de <i>DRIVE</i> . Resultados del experimento de la Sec. 5.3.1. . .	97

A.4. Representación de la predicción de los vasos sanguíneos (píxeles rojos) con las características de las diferentes variantes del operador <i>LBP</i> a múltiples escalas en una imagen de validación de <i>DRIVE</i> . Resultados del experimento de la Sec. 5.3.2. . . . .	98
A.5. Representación de la predicción de los vasos sanguíneos (píxeles rojos) con la inclusión de la característica correspondiente a la intensidad en una imagen de validación de <i>DRIVE</i> . Resultados del experimento de la Sec. 5.3.3. . . . .	99
A.6. Representación de la predicción de los vasos sanguíneos (píxeles rojos) con la inclusión de las características de la resolución 2 : 1 en una imagen de validación de <i>DRIVE</i> . Resultados del experimento de la Sec. 5.3.4. . . . .	100

# Índice de tablas

5.1. Métricas de evaluación de los resultados de la pirámide de imagen con la aplicación del operador <i>LBP</i> . Como ejemplo, <i>Núm. total de escalas</i> = 4 indica la utilización simultánea de las características de las escalas 1 : 1, 1 : 2, 1 : 4 y 1 : 8. . . . .	57
5.2. Métricas de evaluación de los resultados de la aplicación de las diferentes variantes del operador <i>LBP</i> a múltiples escalas. . . . .	59
5.3. Métricas de evaluación tras la inclusión de la característica de intensidad de las imágenes con el preprocesamiento aplicado. . . . .	61
5.4. Métricas de evaluación de los resultados de la aplicación de la característica adicional correspondiente a la escala 2 : 1. . . . .	62
5.5. Resultados de la evaluación mediante la métrica <i>F1</i> de las iteraciones de selección de características por validación cruzada. . . . .	64
5.6. Métricas de evaluación sobre el conjunto de validación del dataset <i>DRIVE</i> basada en un filtrado final del conjunto de características <i>LBP</i> . . . . .	65
5.7. Métricas de evaluación sobre el conjunto de validación del dataset <i>DRIVE</i> partiendo del canal verde. . . . .	66
5.8. Métricas de evaluación sobre el conjunto de validación del dataset <i>DRIVE</i> sintonizando automáticamente los parámetros de la etapa de preprocesamiento. . . . .	67
5.9. Métricas de evaluación sobre el conjunto de validación del dataset <i>DRIVE</i> tras la inclusión de las convoluciones. . . . .	69
5.10. Selección de hiperparámetros seleccionada para la base de datos <i>DRIVE</i> mediante la técnica <i>Grid Search</i> . . . . .	72
5.11. Métricas de evaluación sobre el conjunto de validación del dataset <i>DRIVE</i> en la utilización de parámetros optimizados. . . . .	72
5.12. Métricas de evaluación sobre el conjunto de validación del dataset <i>DRIVE</i> tras la etapa de postprocesamiento. . . . .	73
5.13. Recopilación de las métricas obtenidas en la literatura. Se han resaltado en negrita aquellos valores en los que la evaluación se ha llevado a cabo en las mismas condiciones que en este trabajo. . . . .	76
5.14. Métricas obtenidas en la evaluación de cada una de las imágenes de test de <i>DRIVE</i> . . . . .	77
5.15. Matriz de confusión resultante en la evaluación del conjunto de datos de test de <i>DRIVE</i> . . . . .	77
5.16. Métricas de la consideración global en la evaluación de todos los píxeles de imágenes de test de <i>DRIVE</i> . . . . .	77

5.17. Métricas obtenidas en la evaluación de cada una de las imágenes de test de <i>CHASE</i> . . . . .	83
5.18. Matriz de confusión resultante en la evaluación del conjunto de datos de test de <i>CHASE</i> . . . . .	83
5.19. Métricas de la consideración global en la evaluación de todos los píxeles de imágenes de test de <i>CHASE</i> . . . . .	83
5.20. Métricas obtenidas en la evaluación de cada una de las imágenes de test de <i>STARE</i> . . . . .	86
5.21. Métricas medias de las imágenes de test de <i>STARE</i> . Los valores medios han sido calculados por medio de cada uno de los valores individuales de la Tabla 5.20. . . . .	86

# Acrónimos

<i>CHASE</i>	Child Heart and Health Study in England
<i>CLAFE</i>	Enhance Local Contrast
<i>DRIVE</i>	Digital Retinal Images for Vessel Extraction
<i>GPU</i>	Graphics Processing Unit
<i>LBP</i>	Linear Binary Pattern
<i>LBP NRI</i>	Non Rotation Invariant Linear Binary Pattern
<i>LBP RI</i>	Rotation Invariant Linear Binary Pattern
<i>LBP RIU2</i>	Linear Binary Pattern with Improved Rotation Invariance with Uniform Patterns
<i>LBP VAR</i>	Linear Binary Pattern with Variance Measures of the Contrast
<i>LHP</i>	Local Haar Pattern
<i>LightGBM</i>	Light Gradient Boosting Machine
<i>RGB</i>	Red, Green, Blue
<i>SPM</i>	Spacial Pyramid Matching
<i>STARE</i>	STructured Analysis of the Retina
<i>SURF</i>	Speeded Up Robust Features
<i>SVM</i>	Support Vector Machine
<i>XGBoost</i>	Extreme Gradient Boosting



# 1. Introducción

En este documento, que constituye el *Trabajo de Fin de Máster* del *Máster Universitario en Investigación en Inteligencia Artificial*, se describe en detalle la investigación llevada a cabo para definir una metodología de segmentación de imágenes mediante la utilización de la información proporcionada por el operador *Linear Binary Pattern* basado en *pirámide de imagen*.

Esencialmente, para definir la metodología propuesta se hace uso de dos técnicas diferentes. Por un lado, *Linear Binary Pattern (LBP)*, un operador de procesamiento de imágenes que permite obtener un conjunto de características relevantes de los píxeles de la imagen mediante la descripción de sus vecindades. Concretamente, se usa aquí su variante *pirámide de imagen*, que hace frente a la principal debilidad del operador, es decir, la explosión combinatoria de valores *LBP* al emplear radios y parámetros de vecindad elevados. Y, por otro lado, se usará el clasificador *Light Gradient Boosting Machine* para, mediante las características generadas por medio del operador *LBP*, entrenar un modelo que facilite la segmentación de imágenes a nivel de píxel. Una vez obtenida la clasificación de cada píxel, el proceso de segmentación resulta inmediato.

Con todo ello, se espera generar un procedimiento alternativo de segmentación de imágenes que sea aplicable a diferentes campos, y particularmente el de la medicina, ya que la investigación será abordada por medio de la utilización de bases de datos de imágenes de fondo de ojo con la finalidad de segmentar los vasos sanguíneos. Sin embargo, este trabajo se encontrará centrado en establecer una metodología genérica para la segmentación de imágenes, y, en consecuencia, se realizará una automatización de todos los procesos implicados de manera que pueda ser empleada a otros tipos de bases de imágenes.

Entre los puntos que se abordarán este capítulo se encuentran la motivación que da lugar a la realización de la investigación, es decir, la concreción de las debilidades que se desean fortalecer en el campo de segmentación de imágenes; el problema que se plantea para llevar a cabo el diseño de una nueva metodología; las limitaciones de las que dispone este trabajo de investigación; la hipótesis principal de la investigación y los objetivos que se desean alcanzar; y, finalmente, se detalla la estructura de este documento.

## 1.1. Motivación

La amplia mayoría de estudios realizados en la actualidad en el ámbito de la clasificación de píxeles de una imagen, es decir, la segmentación de estas con objeto de extraer las regiones de interés con un detalle a nivel de píxel, se encuentran centralizados en torno a la utilización de redes neuronales convolucionales. El uso de este tipo de redes presenta una gran cantidad de ventajas [3], entre las que se encuentran la amplia bibliografía relativa al campo, una elevada precisión en los resultados, y la aplicabilidad a diferentes problemas, pero, sin embargo, requiere así mismo un gran esfuerzo computacional para su entrenamiento, así como bases de datos de gran tamaño y una elevada intervención humana en los procesos de definición de la arquitectura y selección de hiperparámetros, entre otras desventajas. Por estas razones se considera de interés el análisis de aproximaciones alternativas que puedan abordar con éxito el problema de clasificación de píxeles sin técnicas de *Deep Learning*.

En la investigación realizada en este trabajo, se desea proporcionar una metodología alternativa, que precise de tamaños de datasets más pequeños que los usados normalmente con técnicas de *Deep Learning*, y, en consecuencia, una menor potencia computacional. Concretamente, aquí se propone el uso del operador *Linear Binary Pattern (LBP)* como principal generador de las características discriminantes necesarias para el entrenamiento de un clasificador tradicional, entendiéndose por «tradicional» cualquier clasificador que no use técnicas de *Deep Learning*. El operador *LBP* ha sido ampliamente utilizado en la literatura con la finalidad de clasificación de texturas gracias a sus propiedades ya que facilita la descripción de las vecindades de cada uno de los píxeles únicamente a partir de los valores de intensidad de la imagen.

Con ello, se pretende definir una metodología que potencie la descripción de cada uno de los píxeles de las imágenes proporcionada por dicho operador, haciendo frente a las debilidades de este en su utilización a diferentes radios y complementándola mediante el uso de una serie de técnicas adicionales como la convolución, de manera que mediante dichas descripciones sea posible llevar a cabo la segmentación de la imagen. La utilización del operador *LBP* a múltiples radios, es decir, con radios superiores a la unidad de manera simultánea, permite clasificar la vecindad de cada píxel de una manera más detallada, en la que radios de mayor valor codifican información de una región más grande de la imagen. La principal debilidad en este proceso es la explosión combinatoria de posibles valores *LBP* que surgen a medida que se utilizan radios y tamaños de vecindad cada vez mayores, creciendo de manera exponencial con los mismos en el proceso de codificación. Los valores *LBP* son etiquetas numéricas que carecen de orden entre sí, y, por tanto, la explosión combinatoria de posibles valores da lugar a un número muy elevado de categorías que no pueden ser propiamente manejadas como números. A modo de ejemplo, el número de posibles valores *LBP* en su utilización con radio unitario y parámetro de vecindad de valor 8 es de  $2^8$ , mientras que un radio 3 con valor de vecindad 24 proporciona un total de  $2^{24}$  posibles valores. Se propone, por lo tanto, la utilización

del operador con radio unitario sobre en una pirámide de imagen, bajo la hipótesis de que el operador  $LBP$  de radio  $n + 1$ , con  $n \geq 0$ , debe ser equivalente al operador  $LBP$  de radio unitario aplicado a la escala  $1 : 2^n$ .

Básicamente, la metodología seleccionada propone una solución al principal problema que posee el operador  $LBP$ : la explosión combinatoria de valores del operador cuando aumenta el radio y el tamaño de vecindad escogidos para aplicarlo. Al tratarse de un operador basado en la comparación de intensidades de píxeles con otros próximos, permite un buen reconocimiento de bordes y puntos intermedios entre objetos de una imagen. Sin embargo, no resulta de utilidad cuando se comparan píxeles pertenecientes a un mismo objeto donde existe una continuidad. Este inconveniente podría solventarse mediante la comparación con píxeles más alejados, o lo que es equivalente, la utilización de un mayor radio en el operador, pero, sin embargo, esta práctica deriva en la explosión combinatoria expuesta.

## 1.2. Definición del problema

Dada la motivación, se plantea como problema a analizar durante la investigación de este trabajo la definición de un conjunto de características discriminantes, que, empleando como medio principal el operador  $LBP$  basado en pirámide de imagen, sean capaces de clasificar cada uno de los píxeles de una imagen con el fin de segmentarla sin aplicar técnicas de *Deep Learning*.

Concretamente, el problema de aplicación elegido para la evaluación de la metodología propuesta consiste en la segmentación de los vasos sanguíneos en imágenes de fondo de ojo. Para ello se emplean tres bases de datos diferentes compuestas por un total de 68 retinografías en las que los vasos sanguíneos han sido manualmente etiquetados. Una de dichas bases de datos ha sido la empleada para la experimentación, sintonización y validación del proceso óptimo de segmentación de las imágenes, evaluando posteriormente las prestaciones de la metodología finalmente desarrollada en las tres bases de datos. Se trata así de analizar también su robustez en bases de imágenes que tengan propiedades ligeramente diferentes a la empleada inicialmente en el proceso de experimentación.

Se pretende realizar la clasificación binaria de los píxeles de retinografías por medio de la información del operador  $LBP$ . Para ello, se generan características utilizando una *pirámide de imagen* en la que se aplica el operador  $LBP$  con radio unitario y vecindad  $P = 8$  en cada nivel, simulando la utilización de radios y parámetros de vecindad mayores, tanto más grandes cuanto mayor sea el nivel de pirámide en el que se aplica.

Se propone la independización del proceso del criterio experto y de la intervención humana automatizando en la medida de lo posible todas las etapas del algoritmo de segmentación de forma que cada una de ellas pueda ser aplicada de manera individual a las diferentes bases de datos, adaptándose a sus propiedades, y generando

un clasificador con capacidad de discriminar los píxeles. Entre las tareas que se automatizarán se encuentran la definición de una etapa de preprocesamiento óptima, previa a la aplicación del operador *LBP*, y la selección de la combinación de parámetros del clasificador.

Así mismo, se llevará a cabo el análisis de características que puedan incrementar las prestaciones de la metodología definida mediante un aumento del poder discriminatorio del modelo. Para ello, se comprobará la utilidad de la aplicación de diferentes técnicas de forma complementaria a las características generadas por medio del operador *LBP*, como la inclusión directa de los valores de intensidad de los píxeles de las imágenes, la utilización de convoluciones, o la aplicación de un postprocesamiento.

Finalmente, se debe considerar que las bases de imágenes escogidas han sido retinografías, las cuales tienen aplicabilidad en el campo de la medicina, donde la correcta identificación de los vasos sanguíneos facilita el diagnóstico de enfermedades dado que se trata de la única parte del cuerpo en la que se pueden observar directamente los vasos sanguíneos, existiendo, por lo tanto, aplicabilidad práctica del problema seleccionado. Entre dichas enfermedades se encuentran [4], por citar algunos ejemplos, la retinopatía hipertensiva, la retinopatía diabética, la vasculitis o el glaucoma.

### 1.3. Limitaciones

Establecidas las motivaciones y definido el problema a abordar durante la investigación realizada, dado que el *Trabajo de Fin de Máster* posee un alcance limitado, y a pesar de que la definición de la metodología propuesta es genérica, es decir, aplicable a diferentes problemas de segmentación de imágenes, únicamente se aplicará a bases de datos de imágenes de retina, realizando la segmentación de los vasos sanguíneos y la evaluación de dicha metodología sobre los resultados proporcionados.

### 1.4. Hipótesis y objetivos

Las hipótesis principales sobre las que se centra la investigación, y que por lo tanto se estudiarán y contrastarán durante la misma, son las siguientes:

*Equivalencia entre la aplicación del operador Linear Binary Pattern a diferentes radios con respecto a la aplicación de dicho operador con radio unitario en una pirámide de la imagen original, generando una descripción del entorno de cada uno de los píxeles más detallada que mediante su clasificación individualizada permita la segmentación de imágenes. De una manera más formal, dada una pirámide de imagen, se plantea la hipótesis de que el operador LBP de radio  $n + 1$  con  $n \geq 0$ , debe ser equivalente al operador LBP de radio 1, pero aplicado al nivel de pirámide  $1/2^n$ .*

*Posibilidad de obtención de clasificadores tradicionales competitivos entrenados con las características generadas por el operador LBP basado en pirámide de imagen. Entendiendo aquí por "clasificador tradicional" aquel que no está basado en técnicas de Deep Learning.*

Para intentar validar ambas hipótesis se plantean los siguientes objetivos:

- Elección de un clasificador tradicional que no esté basado en técnicas de *Deep Learning* y que use como características las proporcionadas por el operador *LBP* aplicado a pirámide de imagen. Establecimiento, por lo tanto, de una metodología que permita realizar la clasificación individual de los píxeles de una imagen para, posteriormente, facilitar la realización de un proceso de segmentación de los objetos de interés. Con ello se pretende prescindir de la utilización de un hardware específico como *GPUs* y disminuir considerablemente el tiempo de entrenamiento sin recurrir a bases de datos de gran tamaño.
- Diseño e implementación de una etapa de preprocesamiento automático que adapte la metodología propuesta a las particularidades de cada base de imágenes.
- Evaluación del clasificador en el problema de segmentación de vasos en imágenes de retina. Dicha evaluación se realizará en tres bases de datos de retinografías diferentes.
- Estudio de técnicas de postprocesamiento para tratar de mejorar la segmentación obtenida con el etiquetado del clasificador.

## 1.5. Estructura del documento

La estructura del resto del documento se indica a continuación. En el capítulo 2 se encuentra el análisis detallado del *Estado de la Cuestión*, el cual está dividido en las técnicas de segmentación tradicionales y las de *Deep Learning*, para facilitar la comparación de los resultados de este trabajo con la literatura más moderna y con los trabajos con mayor similitud al desarrollado. A continuación, en el capítulo 3, se realiza una descripción de las principales técnicas empleadas con la finalidad de facilitar al lector aquellos conocimientos básicos que puedan ser necesarios para la comprensión del trabajo de investigación. El capítulo 4 contiene la descripción propia de la investigación y la metodología finalmente obtenida. Los resultados generados por medio de la metodología definida se muestran en el capítulo 5. Por último, en el capítulo 6 se incluyen las conclusiones del estudio realizado, así como el trabajo futuro que puede derivar de la investigación realizada.



## 2. Estado de la Cuestión

En este capítulo se proporciona información acerca de investigaciones similares, tanto en las que se hayan empleado hipótesis equivalentes a las de esta investigación, como estudios en los que se haya resuelto el mismo problema de segmentación en vasos sanguíneos para facilitar una comparación de los resultados de este trabajo. Se ha llevado a cabo, por lo tanto, un análisis del *Estado de la Cuestión*, revisando una gran variedad artículos y realizando una selección de los más relevantes de cara a disponer de referencias en este trabajo de investigación.

Para la búsqueda de literatura relevante se ha dado prioridad a la afinidad de los artículos con metodologías similares a la escogida para este trabajo, seleccionando, en la medida de lo posible, artículos recientes de revistas. Así mismo, la selección de bases de imágenes de retinografías a emplear en este trabajo se verá influenciada por las escogidas en los estudios que se analicen en este capítulo con la finalidad de facilitar una comparación posterior de los resultados de evaluación.

Por lo tanto, el análisis de la literatura se ha basado principalmente en las siguientes líneas:

- Comprobación de la existencia de metodologías similares a las del problema de este trabajo de investigación.
- Análisis de los resultados obtenidos en estudios que por medio de herramientas y metodologías diferentes realizan la tarea de segmentación de vasos sanguíneos en imágenes de fondo de ojo, es decir, el mismo problema de aplicación que en este trabajo.

### 2.1. Trabajos relacionados con el operador *LBP* y la pirámide de imagen

A continuación, se detalla una serie de investigaciones con metodología similar a la que se han definido para la resolución del problema planteado en este trabajo, es decir, aquellas que combinan la utilización del operador *LBP* con la pirámide de imagen de forma equivalente al empleo del operador con radios superiores. La principal diferencia de las investigaciones de estas características con el problema propuesto es la aplicación escogida, dado que se centran en la clasificación de imágenes en lugar de su segmentación. La clasificación de imágenes, y, concretamente

de texturas, es el uso más extendido en la literatura para el operador *LBP*. Por lo tanto, a pesar de que únicamente los dos primeros artículos emplean los histogramas generados por el operador, prácticamente toda la literatura lleva a cabo una clasificación de la imagen en su totalidad en lugar de una clasificación a nivel de píxel como se realizará en este trabajo.

Los criterios de selección de los artículos son los indicados en el apartado previo, es decir, el empleo de técnicas equivalentes a las escogidas, preferentemente publicados en revistas y lo más recientes posible.

- ***Face Recognition with Local Binary Patterns, Spatial Pyramid Histograms and Naive Bayes Nearest Neighbor classification*** [5]. Se introducen dos algoritmos cuyo objetivo es realizar tareas de reconocimiento facial. En cuanto a este campo, el reconocimiento facial, existe una principal carencia de los algoritmos de detección, que es la necesidad de que los rostros presentes en las imágenes se encuentren alineados en las distintas imágenes que se consideran. La aportación que se realiza en el campo es una mayor flexibilidad en cuanto a esta problemática, para lo que se proponen dos algoritmos que consideran, por separado, herramientas del presente trabajo.

En ambos algoritmos se utiliza el operador *Linear Binary Pattern (LBP)* debido a su simplicidad y eficiencia computacional para proporcionar buenos resultados en el reconocimiento facial. De forma previa a la aplicación del operador, se dividen las imágenes en múltiples regiones de iguales dimensiones, de las que se comparan los histogramas *LBP* obtenidos en las posiciones homólogas de las imágenes. Se concatenan los correspondientes vectores y se realiza la comparación de las imágenes mediante el vector global concatenado. La principal desventaja de este procedimiento es la consideración de que las mismas regiones representan los mismos elementos en las imágenes comparadas, es decir, las mismas partes del rostro. Como problema adicional se tiene que, si se aumenta el tamaño de estas regiones, se mejora el comportamiento para situaciones en que las caras no se encuentren alineadas, pero se disminuye el poder de discriminación entre los múltiples rostros.

En adición al operador *LBP*, las dos versiones indicadas del algoritmo se complementan con las siguientes herramientas:

- ***Spatial Pyramid Matching (SPM)***: con la finalidad determinar las dimensiones óptimas de las regiones en que se dividen las imágenes, se lleva a cabo el cálculo de los histogramas *LBP* de la misma imagen a diferentes resoluciones. En primer lugar, se emplea un histograma sobre la imagen completa, posteriormente la imagen se divide en cuatro regiones iguales, calculando un histograma para cada región, y así sucesivamente para un número definido de niveles. En el artículo, para la comparación de los histogramas se otorga un peso menor a las resoluciones bajas ya que es menos probable que representen una misma parte de la cara.

- *Naïve-Bayes Nearest Neighbor*: en lugar de concatenar los vectores correspondientes a cada histograma, se conservan por separado, almacenando adicionalmente en dichos vectores la información de la posición espacial en coordenadas del centro de la correspondiente región. Los rostros se comparan mediante el conjunto total de vectores. Para la clasificación se emplea un *Nearest Neighbor*, junto con la conjetura de *Naïve-Bayes* de independencia entre los atributos, es decir, los vectores, y la clase, el sujeto cuyo rostro se está identificando. Por lo tanto, se puede asumir que la probabilidad de que la imagen pertenezca al sujeto es igual al producto de probabilidades de que cada vector pertenezca al sujeto, estimando posteriormente la probabilidad final mediante el vecino más cercano. La principal diferencia con métodos previos es que no comparan imágenes entre sí, sino que se compara la totalidad de los vectores correspondientes a las imágenes de cada sujeto.

Para el estudio de los resultados se emplean distintas bases de datos, con diferente número de sujetos, variación de la pose de las caras, expresión o particularidades como el uso de gafas. Se comparan los resultados obtenidos para los dos algoritmos diseñados y otros existentes previamente. Los resultados se ofrecen en forma de porcentajes de acierto en la clasificación, de los que se concluye que el algoritmo que mejor se comporta ante la variación de las poses y orientaciones es *Naïve-Bayes Nearest Neighbor*, pero, sin embargo, en comparación con los métodos previos, el *Spatial Pyramid Matching* también ofrece mejoras relevantes.

Además de la utilización del operador *LBP*, es posible encontrar en el artículo dos herramientas escogidas para la metodología del *Trabajo Final de Máster*, el empleo del operador sobre diferentes escalas de la imagen, tal y como expone la principal hipótesis, concluyendo en los resultados que se obtienen mejoras relevantes en el rendimiento del operador en su utilización, y el empleo de un modelo tradicional, pero que, sin embargo, emplea técnicas de *clustering* y por lo tanto no proporciona una relación directa con el análisis a nivel de píxel.

- ***Uniform Local Binary Pattern for Fingerprint Liveness Detection in the Gaussian Pyramid*** [6]. El artículo se encuentra en la línea de las investigaciones en las que se plantea el uso de la técnica de pirámide de imagen como complemento a *Linear Binary Pattern*. La razón por la que se ha seleccionado este artículo es el nivel de detalle sobre los fundamentos del algoritmo que se emplea y también el hecho de que al ser un estudio reciente permite conocer las técnicas más modernas en el área.

Se propone un método de detección de huellas dactilares falsas para mejorar la eficacia que tienen estos sensores en los dispositivos electrónicos. En lugar de emplear el algoritmo tradicional *LBP*, se plantea una variante denominada *Uniform Linear Binary Pattern* (véase la Sec. 3.1.2) junto con la técnica conocida como *Spatial Pyramid*, también empleada en el estudio [5] previamente

analizado. Como clasificador se utiliza *Support Vector Machine (SVM)*, con dos posibles clasificaciones, como huella real o falsa.

Se emplea un preprocesamiento conocido como *Gaussian Pyramid Filter*, donde se seleccionan grupos de píxeles, que son sucesivamente cada vez más pequeños, con el objetivo de desenfocar la imagen, lo que facilita la detección de bordes u objetos característicos. De forma concreta, el método Gaussiano previene que se forme el efecto de *aliasing* en la imagen, para ello, suaviza los datos de la imagen y los extiende a los píxeles vecinos de forma previa a la selección de píxeles, con lo que no se borra completamente la información de los que son descartados.

Se trata de un artículo muy completo en el que se indican de manera teórica los distintos procedimientos llevados a cabo en el preprocesamiento de las imágenes. Así mismo, se proporciona información de los conjuntos de datos empleados y los resultados obtenidos, logrando una mejora respecto a los proporcionados por otras investigaciones previas.

De nuevo, se tienen diferencias con la metodología escogida para este trabajo, ya que la pirámide se genera mediante un filtro gaussiano, en el que se facilita la identificación de los bordes, pero, sin embargo, se pierde información a nivel individual de los píxeles, y por lo tanto no se corresponde exactamente con el objetivo escogido de representación de una mayor cantidad de información de la vecindad de los píxeles simulando un mayor radio de aplicación del operador. Así mismo, se difiere en la selección y empleo del clasificador, el cual se utiliza a nivel de imagen por medio de los histogramas en lugar de su aplicación directa a nivel de píxel para su clasificación individual, ya que se tiene un problema de clasificación de imágenes en lugar de segmentación de las mismas.

- ***Topology-based image segmentation using LBP pyramids*** [7]. La aplicación escogida en el estudio es una de las más comunes del operador *LBP*, la clasificación de texturas, buscando optimizar la información proporcionada por el operador para facilitar un aumento del poder predictivo de los datos de partida a la hora de identificar las diferentes texturas.

Se detalla un complejo algoritmo que mediante la combinación de la información que proporciona la codificación de la información resultante de la comparación de intensidades junto con la obtenida por la pirámide de imagen, mediante la fusión de las regiones próximas con características similares y la identificación de valores máximos, mínimos, pendientes, y otras propiedades de las regiones, es capaz de segmentar elementos presentes en la imagen, es decir, regiones con propiedades comunes, generando por lo tanto una descripción más detallada de las texturas presentes.

Dado que la metodología difiere en gran medida con la escogida, no se proporciona en este apartado un análisis más detallado del artículo, pero, sin embargo, cabe destacar que también emplea el operador *LBP* aplicado a pirámide de imagen, y, que al igual que este *Trabajo Fin de Máster*, se encuentra

orientado a la segmentación de imágenes, demostrando las ventajas del uso del reescalado para disponer de información de píxeles más alejados, sin necesidad de aumentar el radio del operador *LBP*. Así mismo, al igual que con la metodología establecida para el presente estudio, se prescinde de la utilización de los histogramas generados por el operador, empleando directamente la información individual de los diferentes píxeles, ya que con ello no se conserva la información de la posición espacial de los objetos segmentados.

- ***Gaussian image pyramid based texture features for classification of microscopic images of hardwood species*** [8]. Al igual que en el artículo previamente comentado, el objetivo de la investigación es la clasificación de texturas, concretamente para la clasificación de maderas, donde se obtienen resultados considerablemente más precisos que las técnicas tradicionalmente empleadas.

En el artículo, además del operador *LBP*, se emplean otros operadores que realizan la extracción de información de las imágenes como *Local Configuration Pattern*, que constituye una ampliación del operador *LBP* que mediante la comprobación de la relación entre parejas de píxeles [9] proporciona información de la configuración microscópica de la textura, y *Local Phase Quantization*, descriptor para clasificación de texturas con especial capacidad para el análisis de imágenes difuminadas que emplea una estimación de la *transformada de Fourier* a nivel local de la convolución que sería equivalente a la aplicación de un filtro que diese lugar a la difuminación que exista en la imagen [10]. Dichos algoritmos se combinan con un *Gaussian Pyramid Filter* con siete niveles de descomposición para mejorar la precisión de la posterior clasificación. El clasificador utilizado es un *SVM* de comportamiento lineal.

Los pasos seguidos por el algoritmo generado son los siguientes:

1. Aplicación de un preprocesado que transforma la imagen *RGB* en escala de gris. Descomposición de la imagen mediante el filtro gaussiano aplicado a diferentes escalas.
2. Extracción de vectores de características en las diferentes escalas sobre las que se ha aplicado el filtro gaussiano. Los vectores son escalados entre valores pertenecientes al intervalo  $[0, 1]$ .
3. Aplicación del clasificador *SVM*.

Aunque la clasificación de texturas no es el objeto del *Trabajo de Fin de Máster*, y, por lo tanto, sus resultados no son equiparables a los que se obtengan de esta investigación, es de interés el conocimiento de las técnicas empleadas para la extracción de información de la imagen en investigaciones basadas en la utilización del operador *LBP* a diferentes escalas.

## 2.2. Análisis de trabajos relacionados con la segmentación de vasos en retinografías

Con la finalidad de analizar la literatura de investigaciones relativas a la segmentación de vasos sanguíneos, se proporcionan a continuación las principales características de algunos estudios. De esta forma será posible establecer una referencia de los resultados de evaluación de investigaciones recientes que permita realizar una comparación con los futuros resultados de este trabajo.

Se ha llevado a cabo una selección de artículos recientes que comprenden trabajos tanto supervisados como no supervisados, con una subdivisión entre los supervisados que emplean redes neuronales convolucionales y los que se basan en clasificadores tradicionales, siendo estos últimos los de mayor similitud con respecto a este trabajo de investigación, pero entre los que se encuentra una literatura más acotada.

### 2.2.1. Técnicas de aprendizaje no supervisado

En esta sección se recogen trabajos cuya aproximación se basa en metodologías consistentes en filtros o técnicas de aprendizaje automático sin la utilización de la etiqueta real de los píxeles. Por lo general, este tipo de metodologías proporcionan una menor precisión que aquellas en las que se realiza un aprendizaje supervisado. La etiqueta real únicamente se emplea en la fase final de evaluación, prescindiendo de ella para un proceso de entrenamiento. A continuación, se detallan dos metodologías que se han seleccionado debido a la evaluación de los resultados respecto a las bases de datos más comunes entre la literatura analizada también en los trabajos supervisados (véase la Sec. 2.2.2). Así mismo, dado que la literatura es limitada, otro criterio empleado en su selección es que se trata de aquellas con fecha de publicación más reciente para que puedan ser empleadas en la medida de lo posible como referencia.

- *Automatic Extraction of Retinal Blood Vessel Based on Matched Filtering and Local Entropy Thresholding* (2015) [11]. La metodología escogida comienza por una corrección de la iluminación no uniforme junto con el método *CLAHE* para mejorar el contraste, y posteriormente se mejora la imagen mediante un núcleo gaussiano bidimensional con diferentes valores del parámetro  $\sigma$  para cubrir las estructuras gruesas y finas de los vasos sanguíneos de la retina, es decir, un filtro de emparejamiento. Para identificar automáticamente los vasos finos o gruesos se emplea un clasificador de vector soporte, empleado para calcular el hiperplano óptimo. Como criterio final, se utiliza un umbral de entropía local por medio de la intensidad de gris para la extracción final de los vasos sanguíneos.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	-	0,7481	0,9748	0,9597
CHASE	-	-	-	-
STARE	-	0,7298	0,9831	0,9532

A diferencia de lo comúnmente realizado en los trabajos que emplean estas bases de datos, se han utilizado las etiquetas generadas por el segundo observador.

- *Retinal vessel segmentation using a finite element based binary level set method* (2017) [12]. Se propone una combinación de técnicas de etiquetado de vasos sanguíneos mediante la modificación de filtros de emparejamiento tradicionales, utilizando un método basado en elementos finitos que proporciona información más detallada acerca de los bordes, reduciendo de esta forma la complejidad de selección de un valor límite. En los filtros de emparejamiento una función Gaussiana puede ser empleada para determinar las secciones entre vasos sanguíneos, sin embargo, un valor límite global no proporciona resultados lo suficientemente precisos.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	-	0,784	0,979	0,951
CHASE	-	-	-	-
STARE	-	0,753	0,981	0,952

Las métricas se proporcionan empleando como referencia al primero de los observadores manuales en ambas bases de datos.

La evaluación de los resultados obtenidos por ambos trabajos ha sido realizada sobre el conjunto total de 20 imágenes de las bases de datos sobre las que se proporcionan los resultados, siendo las 20 correspondientes al conjunto de test para el dataset *DRIVE* y las 20 totales del dataset *STARE*. No se indica si en el cálculo de las métricas se descartan los píxeles que no forman parte propiamente del fondo del ojo, es decir, si se aplica una máscara que elimine los bordes que no formen parte de las capturas determinando así un *campo de visión*.

## 2.2.2. Técnicas de aprendizaje supervisado

En esta sección se analizan investigaciones realizadas por medio de entrenamientos supervisados, es decir, en los que se ha empleado la etiqueta real de los píxeles como parte del entrenamiento del modelo de clasificación. Se realiza una división entre los relativos a *Deep Learning* y los que emplean clasificadores tradicionales.

### 2.2.2.1. Técnicas basadas en redes neuronales convolucionales

Debido al elevado rendimiento que proporciona la utilización de redes neuronales convolucionales en el proceso de aprendizaje automático supervisado en tratamiento

de imágenes, difícilmente alcanzable mediante otros métodos, y las ventajas expuestas previamente (véase la Sec. 1.1), las investigaciones más recientes de segmentación de vasos sanguíneos en imágenes de fondo de ojo suelen centrarse en estos sistemas de clasificación. En la literatura existe una gran variedad de técnicas de segmentación en torno a las bases de datos de retinografías que proporcionan una solución mediante este tipo de redes, concretamente, es muy común la utilización de la red neuronal convolucional *U-Net*, la cual ha sido contrastada en el campo de clasificación en imágenes de origen médico. De entre los estudios basados en técnicas de *Deep Learning*, se indican a continuación algunas de las más recientes y que han obtenido mejores resultados.

Debe considerarse que este trabajo no pretende competir con técnicas de *Deep Learning*, tal y como se indica en la Sec. 1.4 correspondiente a los objetivos, sino avanzar en la creación de métodos que utilicen menos recursos computacionales, requieran menos datos, tiempos de entrenamiento y prescindan de la necesidad de configuración manual de arquitecturas para redes neuronales, y que, por lo tanto, el análisis se proporciona con la finalidad de dar una visión global del *Estado de la Cuestión* relacionado con el problema de segmentación de vasos en imágenes de retina.

- *Bi-SANet—Bilateral Network with Scale Attention for Retinal Vessel Segmentation* (2021) [13]. La metodología propuesta tiene como objetivo dar solución a problemas que no son correctamente optimizados en la segmentación de vasos sanguíneos en imágenes de fondo de ojo mediante otras redes convolucionales existentes. Concretamente, la segmentación de las zonas con menos contraste y con lesiones tiene un carácter ineficiente en las redes *U-Net* comúnmente utilizadas para este tipo de imágenes, buscando con esta optimización la mejora de la sensibilidad manteniendo la especificidad y precisión. Con ello, se propone una red *U-Net* que trabaja a múltiples escalas.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	0,8382	0,8890	0,9772	0,9693
CHASE	0,8139	0,8371	0,9852	0,9759
STARE	0,8436	0,8290	0,9894	0,9770

- *SA-Net: A scale-attention network for medical image segmentation* (2021) [14]. Se proporciona una modificación de *U-Net* en la que las características son introducidas a diferentes escalas.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	0,8289	0,8252	0,9764	0,9569
CHASE	0,8280	0,8199	0,9827	0,9827
STARE	-	-	-	-

Las anotaciones empleadas son las proporcionadas por el primer observador manual. El campo de visión para el dataset *CHASE* ha sido creado manualmente para filtrar aquellos píxeles que no corresponden propiamente al fondo de ojo.

- *Automatic Retinal Blood Vessel Segmentation Based on Fully Convolutional Neural Networks* (2019) [15]. La investigación se propone como mejora de la tarea de segmentación de vasos sanguíneos en retinografías con respecto a las técnicas existentes en el momento de su desarrollo. Se propone un sistema de segmentación automática de los vasos sanguíneos por medio de redes neuronales convolucionales profundas, que integran métodos de preprocesamiento de los datos.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	0,8321	0,8325	0,9838	0,9706
CHASE	0,8243	0,8453	0,9862	0,9773
STARE	0,8531	0,8522	0,9880	0,9777

Como metodología de evaluación, para el dataset *DRIVE* se emplea la división propuesta por los creadores de la base de imágenes, es decir, un total de 20 imágenes de entrenamiento y 20 de test, en el dataset *STARE* se emplea la metodología «*leave-one-out*», con una evaluación cruzada en la que un conjunto de 19 imágenes son empleadas para el entrenamiento y la restante para evaluación, alternando la imagen de evaluación en las sucesivas iteraciones, y en el dataset *CHASE* se emplea una división en la que las 20 primeras imágenes son empleadas para el entrenamiento y las restantes 8 de test.

### 2.2.2.2. Técnicas de aprendizaje supervisado no basadas en *Deep Learning*

Dado que los métodos incluidos en este apartado se basan en técnicas supervisadas pero sin implementación de *Deep Learning*, proporcionando una mayor relevancia a la definición de una metodología de generación de características que permitan discriminar correctamente los vasos sanguíneos en lugar de centrar los esfuerzos en el diseño de la arquitectura de una red neuronal, se busca el objetivo de alcanzar resultados de evaluación en este trabajo que puedan ser competitivos con los estudios que se detallan en esta sección.

- *A new supervised retinal vessel segmentation method based on robust hybrid features* (2016) [16]. Se combina un total de 17 características obtenidas por diferentes algoritmos de modificación de contraste, intensidad, morfología y otros filtros que proporcionan información individual de cada píxel. El clasificador utilizado es *Random Forest*.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	-	0,7545	0,9801	0,9513
CHASE	-	-	-	-
STARE	-	0,7556	0,9837	0,9605

Al igual que en los restantes trabajos, en el dataset *DRIVE* se emplea la distribución de imágenes sugerida en la documentación de la base de datos, mientras que en el dataset *STARE* se utiliza un 1% del total de los píxeles

a modo de entrenamiento y se evalúan los resultados sobre la totalidad de imágenes, no manteniendo por lo tanto una plena independencia. Se detalla la metodología empleada para crear una máscara correspondiente a las imágenes del dataset *STARE* que permita filtrar el *campo de visión*. Emplea las etiquetas definidas por el primer observador manual.

- *Retinal vessel segmentation in colour fundus images using Extreme Learning Machine* (2017) [17]. Se extrae un conjunto de 39 vectores, que consisten en características locales, rasgos morfológicos, congruencia de fase, hessiana y divergencia de campos vectoriales. El clasificador empleado es *Extreme Learning Machine*, aplicado a nivel de píxel. Posteriormente se aplica un postprocesamiento para la eliminación de ruido.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	-	0,7140	0,9868	0,9607
CHASE	-	-	-	-
STARE	-	-	-	-

Se utiliza la división estándar del dataset *DRIVE*, pero no se indica la etiqueta de qué observador es empleada.

- *An innovate approach for retinal blood vessel segmentation using mixture of supervised and unsupervised methods* (2020) [18]. Se combinan metodologías no supervisadas y supervisadas mediante la aplicación de descriptores de características locales como *LHP* y *mSURF* y la utilización posterior de los clasificadores *Random Forest*, *Support Vector Machine* y *Adaboost*.

Dataset	F1-score	Sensibilidad	especificidad	Precisión
DRIVE	-	0,786	0,973	0,958
CHASE	-	-	-	-
STARE	-	0,831	0,963	0,953

En la evaluación se emplean validaciones cruzadas con un total de 10 particiones, sin indicar cuál de los observadores ha sido empleado.

## 3. Conocimiento previo

En este capítulo se recogen, con la finalidad de facilitar los conocimientos básicos que sean requeridos para la comprensión de los próximos capítulos, las principales herramientas que se utilizarán en la definición de la metodología para el desarrollo de la investigación.

### 3.1. *Linear Binary Pattern (LBP)*

Se trata de un operador [19] que por medio de las intensidades de los píxeles de una imagen genera una descripción de la vecindad de cada píxel a partir de una comparación con los circundantes. Este operador ha sido ampliamente utilizado en la literatura existente en tareas de clasificación de texturas [8], así como, por ejemplo, en la identificación de patrones [5], y en algunas tareas similares a las de este trabajo en segmentación de imágenes [7], aunque su utilización para este campo es reducida.

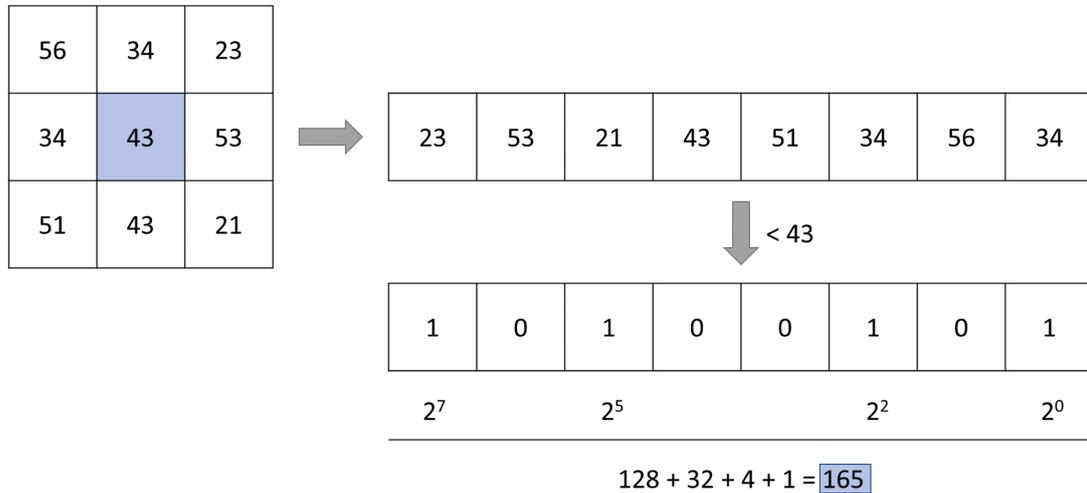
El operador se emplea a nivel de píxel y se define para un radio  $R$  y un tamaño de vecindad  $P$ . A modo de ejemplo, al establecer  $R = 1$  y  $P = 8$  la transformación se realiza considerando un total de 8 píxeles circundantes (véase la Fig. 3.1). Determinada la vecindad por medio de dichos parámetros, se genera un vector en el que cada una de sus componentes representa una comparación de intensidad del píxel vecino con el que está siendo transformado, proporcionando una codificación binaria de cuya transformación a base decimal se obtiene el valor  $LBP$ . Tras la transformación de todos los píxeles se obtiene una imagen de las mismas dimensiones que la imagen original<sup>1</sup> pero en la que los nuevos valores representan únicamente una categoría en la que no existe una relación ordinal a pesar de ser numéricos. En la Fig. 3.1 se observa un ejemplo de cálculo con  $R = 1$  y  $P = 8$ .

Al modificar los valores de radio y tamaño de vecindad se establece la región de píxeles vecinos. Con el ejemplo indicado, compuesto por un total de 8 valores de intensidad, se generan un total de  $2^8 = 256$  categorías para el posible valor resultante tras la codificación decimal. Sin embargo, si se desea realizar una categorización de vecindades de regiones de mayor tamaño, como por ejemplo podría ser con valores  $R = 2$  y  $P = 16$ , se tendría un vector de 16 componentes donde el número de posibles categorías sería  $2^{16} = 65536$ , de donde es posible deducir la explosión combinatoria

---

<sup>1</sup>Para los píxeles situados en los límites de la imagen, en los que no se tienen todos los píxeles vecinos necesarios, generalmente se proporciona un valor de intensidad mínimo o máximo a los píxeles inexistentes, facilitando de esta forma que pueda ser realizada la operación.

de valores *LBP* creciente de manera exponencial con dichos parámetros, expuesta en el Cap. 1.



**Figura 3.1.:** Ejemplo de cálculo del operador *LBP* para un píxel de análisis (píxel central), considerando el caso más simple ( $R=1$  y  $P=8$ ).

Una vez aplicado el operador *LBP* sobre la imagen, generalmente se definen histogramas que indican la frecuencia de aparición de las categorías. Dicho histograma puede ser utilizado para identificar texturas o clasificar imágenes. Sin embargo, dado que el empleo del operador en este trabajo tiene como fin una clasificación a nivel de píxel, se consideran directamente los valores *LBP* de los diferentes píxeles.

### 3.1.1. Limitaciones

El operador *LBP* posee una principal limitación, el número de píxeles vecinos que dan lugar al vector binario. Tal y como se ha indicado en el apartado introductorio de esta sección, la utilización de parámetros de radio y vecindad elevados supone una explosión combinatoria de categorías *LBP*, llegando a ser ineficiente el almacenamiento de tal cantidad de información y, por lo tanto, dificultando enormemente la creación del modelo. Así mismo, al aumentar considerablemente el número de categorías disminuye la frecuencia de aparición individual de cada una de ellas y aumenta la cantidad de datos necesarios para realizar un entrenamiento sin pérdida de poder discriminatorio.

### 3.1.2. Variantes del operador *LBP*

Para algunas aplicaciones la información proporcionada por el operador *LBP* es superior a la necesaria cuando, por ejemplo, se desean identificar patrones como bordes o esquinas, pero la orientación que poseen no es relevante. Dependiendo del problema, un número tan elevado de categorías genera adición de ruido y una simplificación de las mismas puede proporcionar mejores resultados, de donde surgen algunas variantes del operador. Así mismo, existen variantes que proporcionan información adicional como la varianza de intensidades de la región que no son contempladas por el operador original. Entre dichas variantes se encuentran:

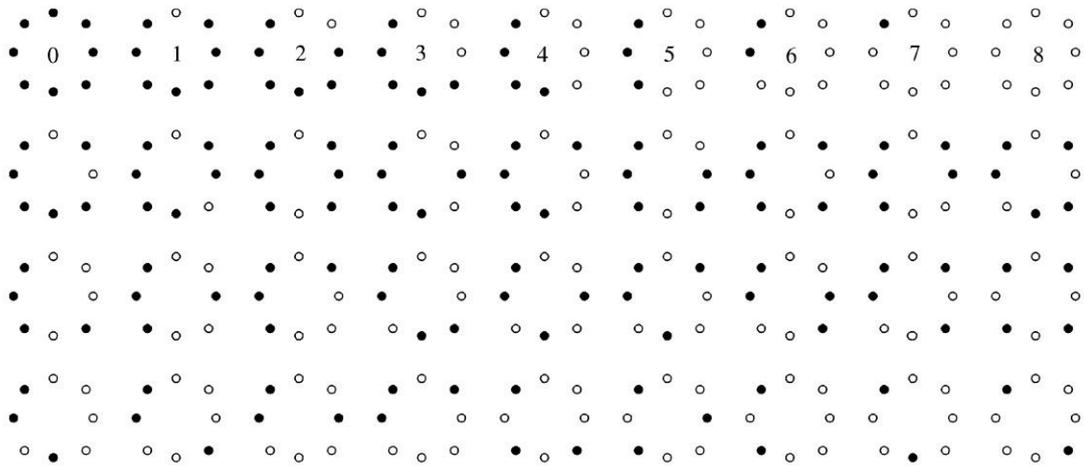
- *Rotation invariant (RI)* [1]: versión invariante a la rotación en la que, calculado el vector en base binaria, se escoge como valor de la categoría *LBP* el menor de los valores en base decimal de los vectores binarios que surgen de las rotaciones de las posiciones de las componentes del vector original. En la Fig. 3.2 se encuentra un ejemplo del paso adicional respecto al operador original que genera esta variante. Para  $R = 1$  y  $P = 8$  esta variante proporciona un total de 36 posibles valores únicos (véase Fig. 3.3).

1	0	1	0	0	1	0	1	$128 + 32 + 4 + 1 = 165$
1	1	0	1	0	0	1	0	$128 + 64 + 16 + 2 = 210$
0	1	1	0	1	0	0	1	$64 + 32 + 8 + 1 = 105$
1	0	1	1	0	1	0	0	$128 + 32 + 16 + 4 = 180$
0	1	0	1	1	0	1	0	$64 + 16 + 8 + 2 = 90$
0	0	1	0	1	1	0	1	$32 + 8 + 4 + 1 = 45$
1	0	0	1	0	1	1	0	$128 + 16 + 4 + 2 = 150$
0	1	0	0	1	0	1	1	$64 + 8 + 2 + 1 = 75$

**Figura 3.2.:** Continuación del ejemplo de la Fig. 3.1 para obtener la versión invariante a la rotación para el caso  $R = 1$  y  $P = 8$ . Las ocho codificaciones binarias mostradas serían equivalentes y 45 el valor finalmente escogido.

- *Improved rotation invariance with uniform patterns (RIU2)* [1]: disminuye el número de posibles categorías o valores del operador *RI* de 36 a 9 (en el caso  $R = 1$  y  $P = 8$ ) mediante la búsqueda de aquellos píxeles con un patrón uniforme en la generación del vector binario. Se realiza una cuenta del número de

transiciones entre valores 0/1, y, si dicha cuenta es superior a 2 se determinará que el patrón no es uniforme y se le asigna la categoría genérica con valor  $v = P + 1$ , siendo  $P$  el tamaño de vecindad. Para los patrones restantes el valor de la categoría es igual al número de *bits* con valor 1. En la Fig. 3.3 se muestran los patrones indicados.



**Figura 3.3.:** Representación [1] de la totalidad de posibles patrones del operador  $RI$  con  $R = 1$  y  $P = 8$ . En la primera fila se muestran los patrones uniformes de la variante  $RIU2$ .

- *Non rotation-invariant uniform patterns (NRI)* [20]: se trata de una variante del operador que únicamente es invariante ante la escala de intensidades de la imagen. Al igual que en la variante  $RIU2$ , se tiene una única categoría para los patrones no uniformes, sin embargo, se realiza la siguiente extensión del número de categorías para los patrones uniformes:
  - Para los patrones uniformes con cero transiciones, es decir, donde todos los *bits* tienen el mismo valor 0/1, se tienen dos categorías en función de dicho valor.
  - Para los patrones uniformes con un total de dos transiciones, cada categoría se corresponde con una de las posibles permutaciones circulares de las posiciones de los píxeles. Por lo tanto, para un radio unitario en el que se tienen 8 píxeles vecinos, el número de posibles categorías correspondientes a este grupo de patrones es de  $8 * (8 - 1) = 56$ .

Por lo tanto, el número total de categorías para  $R = 1$  y  $P = 8$  sería de 59, correspondiente a la suma de la categoría única de patrones no uniformes y las 58 categorías de patrones uniformes, divididas a su vez en las 2 correspondientes a los valores constantes y las 56 categorías en las que existen transiciones de valores 0/1.

- *Rotation invariant variance measures of the contrast of local image texture (VAR)* [21]: variante del operador invariante a la rotación, pero no así a la intensidad. A diferencia de las variantes del operador *LBP* analizadas hasta ahora, en las que la comparación de valores del píxel con respecto a los vecinos es binaria, es decir, únicamente se considera si el valor es superior (o igual) o inferior, en este operador se realiza una medición de la diferencia entre dichos valores. Así mismo, se obtiene un valor numérico cuantitativo continuo en lugar de una categoría cualitativa. La varianza es estimada mediante las siguientes operaciones:

$$s = \sum_i^P I_i; \quad V = \sum_i^P I_i^2; \quad Var = \frac{V - \frac{s^2}{P}}{P}$$

Siendo  $P$  el número de píxeles vecinos e  $I_i$  el valor de intensidad de cada uno de dichos píxeles.

## 3.2. Pirámide de imagen

La técnica *pirámide de imagen* consiste en la aplicación de múltiples iteraciones de reescalado e interpolación sobre la misma imagen para generar imágenes a diferentes resoluciones representando información equivalente a nivel de imagen, pero no así a nivel de píxel. Por lo tanto, como entrada al proceso se tiene la imagen de partida con su resolución original y como salida dicha imagen a múltiples resoluciones, tanto inferiores como superiores, tantas como niveles se le desee proporcionar a la pirámide.

Mediante la combinación de esta técnica con el operador *LBP* se simula la utilización de radios de mayor y menor dimensión de dicho operador sin llevar a cabo una modificación de los parámetros de radio y vecindad. Es decir, se logra la descripción de las propiedades de las vecindades de los píxeles en regiones de diferentes dimensiones mientras que se mantiene un rendimiento equivalente gracias a la disminución drástica del número de categorías en que cada píxel puede ser codificado mediante la utilización del operador *LBP* cuando aumenta el valor de  $R$  y, consecuentemente, el valor de  $P$ .

## 3.3. Clasificador *LightGBM*

La resolución del problema planteado en el trabajo de investigación implica la clasificación de los píxeles de imágenes de fondo de ojo en las que se tiene un *ground truth* que etiqueta cada uno de los píxeles respecto a su pertenencia a vasos sanguíneos, teniendo de esta forma una clasificación binaria. Por lo tanto, la selección del algoritmo de aprendizaje automático empleado para el entrenamiento y posterior predicción ha sido basada en la búsqueda de un clasificador que permita de manera

sencilla la obtención de resultados con una elevada precisión con el menor coste de computación posible.

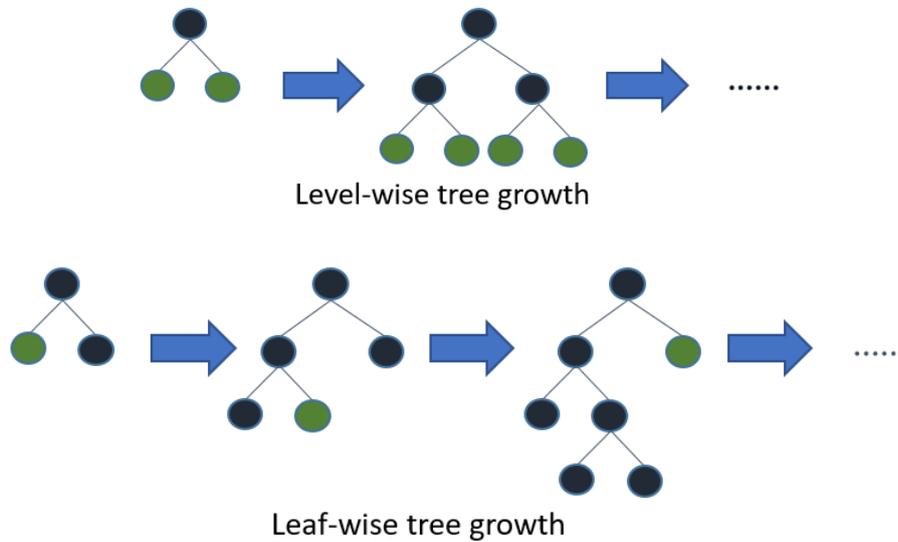
El algoritmo finalmente escogido ha sido *Light Gradient Boosting Machine (LightGBM)*, un clasificador basado en el ensamblado de árboles de decisión mediante la técnica de *boosting*, con unas características muy similares al conocido clasificador *Extreme Gradient Boosting (XGBoost)* [22], pero con un proceso de entrenamiento considerablemente más rápido y que es capaz de proporcionar unos resultados de gran precisión.

El clasificador puede ser explicado mediante la evolución de los clasificadores basados en árboles de decisión:

1. *Árboles de decisión* [23]: modelos de predicción basados en reglas a partir de las cuales se clasifica la información. La finalidad es la creación un modelo capaz de predecir el valor de una variable objetivo a través de reglas de decisión simples inferidas de las características de los datos de entrenamiento. Su construcción se realiza mediante una partición binaria recursiva, que mediante un proceso iterativo divide los datos aplicando una serie de reglas. La división de los datos de entrenamiento en cada una de las bifurcaciones o nodos es realizada mediante la minimización de las desviaciones cuadradas de la media en las nuevas ramas. El proceso de división se realiza en cada una de las ramas generadas a partir de los nodos hasta que se alcanza un tamaño de nodo mínimo, dando lugar a un nodo terminal u hoja del árbol. Mediante dicho tamaño mínimo es posible limitar la capacidad de sobreajuste del árbol a los datos de entrenamiento.
2. *Bagging*: proceso de combinación de los resultados de predicción de modelos contruidos de manera independiente entre sí mediante diferentes submuestras del dataset, entre las que existe repetición (*bootstrapping*), en el que el proceso de clasificación final se realiza mediante un sistema de voto, generando un clasificador que presumiblemente proporciona un mayor poder predictivo que la utilización de los modelos iniciales de forma individual.
3. *Random Forest*: clasificador basado en la técnica de *bagging* en el que los árboles se construyen a partir de subconjuntos diferentes de características de entre todas las disponibles. Es decir, considerando la base de datos estructurada, los diferentes árboles de decisión poseen muestras de los datos tanto a nivel de filas, registros, como de columnas, características, y mediante un posterior sistema de voto se genera una clasificación final.
4. *Boosting* [24]: considerando aprendizaje débil como aquel en el que se genera un clasificador que únicamente es mejor que una estimación aleatoria y aprendizaje fuerte como el que da lugar a un clasificador que con una probabilidad elevada es capaz de generar un modelo que realice una elevada tasa de clasificaciones correctas, la técnica de *boosting* consiste en la mejora de uno o varios algoritmos débiles para dar lugar a uno fuerte con un error más pequeño.

5. *Gradient Boosting* [25]: caso especial de *boosting* en el que los errores son minimizados mediante la optimización conocida como *Descenso del Gradiente*. Se emplean múltiples estimadores, es decir, árboles de decisión, en un proceso secuencial donde los residuos de los errores obtenidos por un estimador son utilizados para actualizar los gradientes que forman parte de la entrada al proceso de entrenamiento del siguiente estimador, siendo el resultado final la combinación de todos los estimadores. Para evitar el sobreajuste, la relevancia de los sucesivos árboles de decisión en el proceso de clasificación final se ve reducida mediante un parámetro denominado *learning rate*.
6. *Extreme Gradient Boosting (XGBoost)*: combinación de técnicas de optimización sobre el *Gradient Boosting* que proporciona mejores resultados en un tiempo de entrenamiento inferior. Entre dichas técnicas de optimización se encuentran:
  - Paralelización del proceso de entrenamiento mediante la construcción de las ramas de cada árbol de manera independiente y simultánea, de forma que, aunque no existe paralelismo entre árboles, sí se tiene en la optimización individual de cada uno de ellos.
  - Regularización en el *Descenso de Gradiente* mediante *Lasso* y *Ridge* [26] para evitar la construcción de modelos excesivamente complejos que den lugar a sobreajustes.
    - La regularización de *Ridge* penaliza la suma de los coeficientes que están siendo ajustados elevados al cuadrado, reduciendo de forma proporcional el valor de todos los coeficientes y con ello la varianza.
    - La regularización de *Lasso* penaliza la suma de los valores absolutos de los coeficientes, forzando a que tiendan a cero aquellos predictores menos relevantes.
7. *LightGBM* [2]: se trata de un algoritmo desarrollado por *Microsoft* con grandes similitudes respecto a *XGBoost*, proporcionando mejoras en términos computacionales respecto a este. La principal diferencia es la construcción vertical en lugar de horizontal de los árboles de decisión, lo que incrementa considerablemente la velocidad del proceso. En la Fig. 3.4 se tienen diagramas de los procesos de construcción de ambos algoritmos, donde se puede observar la construcción vertical e individual para cada una de las ramificaciones de los árboles de decisión en el algoritmo *LightGBM*. De esta forma, el algoritmo realiza el crecimiento de cada rama de manera individual en lugar de considerar todas las hojas de un mismo nivel del árbol simultáneamente. Por lo tanto, las principales ventajas que presenta respecto a *XGBoost* son:
  - Mayor velocidad de entrenamiento.
  - Menor utilización de memoria.
  - Posibilidad de un mayor crecimiento individual de cada una de las ramas.

Sin embargo, para evitar un sobreajuste, es preciso configurar correctamente la profundidad máxima de los árboles.



**Figura 3.4.:** Diagramas de construcción de cada uno de los árboles de decisión en los algoritmos *XGBoost* y *LightGBM*, imágenes superior e inferior respectivamente. Fuente: [2].

### 3.4. Computación evolutiva

Método de computación empleado para realizar procesos de optimización, modelado y simulación en los que imita el comportamiento de la evolución natural, en la que un entorno con unas determinadas características posee una población de individuos que luchan por sobrevivir y reproducirse, y a su vez mutan, favoreciendo, por lo tanto, a los individuos mejor adaptados. Se basa en un método de prueba y error en el que se evalúa la calidad de diferentes soluciones a un problema y aquellas con peor resultado tienden a desecharse.

Entre las principales ventajas de la computación evolutiva se encuentran:

- Aplicabilidad a una amplia gama de problemas.
- Adaptabilidad a problemas muy específicos.
- Obtención de soluciones de calidad en tiempos de computación aceptables.

## 4. Metodología

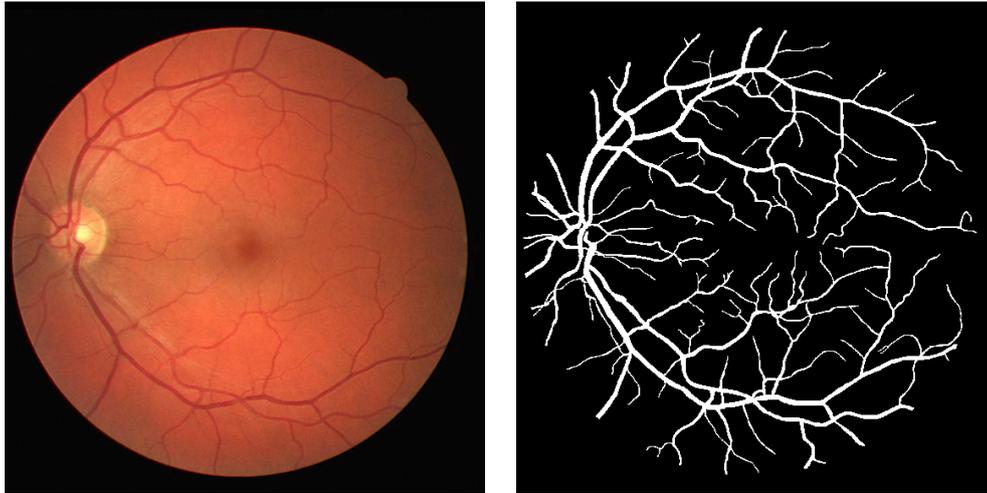
En este capítulo se describe la metodología que ha sido llevada a cabo en el trabajo de investigación de manera que en su conjunto aporte novedades al estado actual de la problemática. De un lado, se describen las bases de datos que se emplearán en los experimentos y, de otro lado, se detalla la metodología necesaria para combinar las herramientas descritas en el capítulo anterior para, de esta forma, poder abordar el problema de investigación definido en este trabajo.

### 4.1. Bases de datos

Como bases de datos con las que desarrollar el proceso de investigación y definición de la metodología que busque alcanzar los objetivos propuestos inicialmente se han seleccionado imágenes de retina. El proceso de segmentación se centra en la identificación automática de los vasos sanguíneos. Para ello se han escogido tres bases de datos que se encuentran públicamente disponibles y que proporcionan la identificación manual por medio de uno o varios expertos de dichos vasos sanguíneos proporcionando imágenes binarias con la etiqueta de cada píxel que pueden ser empleadas a modo de *ground truth*.

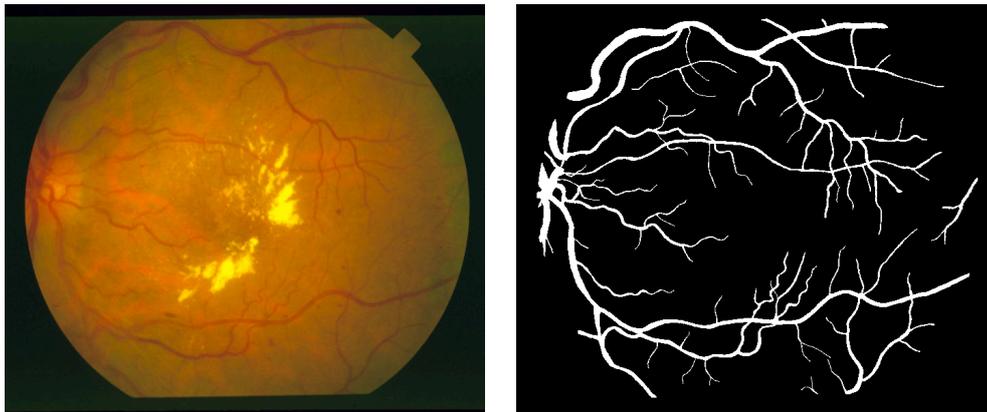
La primera base de datos seleccionada es la conocida como *DRIVE (Digital Retinal Images for Vessel Extraction)* [27, 28], la cual se encuentra compuesta por un total de 40 imágenes con una propuesta de división de 20 como conjunto de entrenamiento y 20 como test. Para el conjunto de test se proporcionan dos etiquetados manuales realizados por diferentes observadores mientras que para el conjunto de entrenamiento únicamente el del primero de ellos. La resolución de las imágenes es de  $565 \times 584$  y se proporciona una máscara conocida como *campo de visión* que permite descartar los píxeles que no forman parte propiamente del fondo del ojo y que, por lo tanto, no deben ser incluidos en el proceso de entrenamiento ni de evaluación. En la Fig. 4.1 se tiene un ejemplo de imagen de la base de datos junto a su *ground truth*.

El segundo conjunto de imágenes empleado recibe el nombre de *STARE (STructured Analysis of the Retina)* [29, 30]. Se encuentra compuesto de un total de 20 imágenes de resolución  $700 \times 605$  y dos sets de imágenes de *ground truth* de diferentes observadores. En esta base de datos no se proporcionan las máscaras que delimitan el *campo de visión* del dispositivo de captura de la imagen. En la Fig. 4.2 se muestra



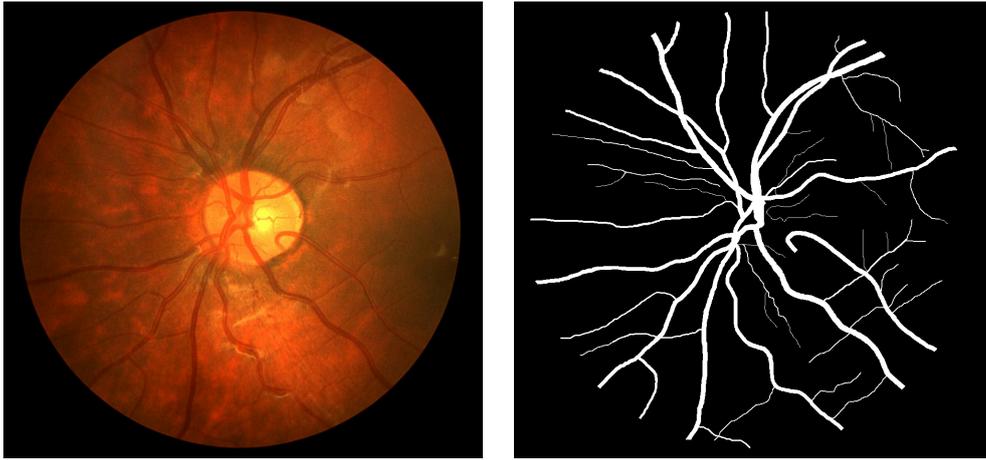
**Figura 4.1.:** Ejemplo de imagen del dataset *DRIVE* junto con su respectivo *ground truth*.

una de las imágenes de la base de datos junto con su correspondiente etiquetado manual del primer observador.



**Figura 4.2.:** Ejemplo de imagen del dataset *STARE* junto con su respectivo *ground truth*.

Como última base de datos se tiene *CHASE (Retinal Image Database)* [31, 32]. Dispone de un total de 28 imágenes a resolución  $999 \times 960$  correspondientes a los dos ojos de 14 individuos. Al igual que *STARE*, cuenta con un etiquetado de vasos de dos observadores para cada una de las retinografías, pero no con el correspondiente *campo de visión*. En la Fig. 4.3 se muestra una imagen del dataset.



**Figura 4.3.:** Ejemplo de imagen del dataset *CHASE* junto con su respectivo *ground truth*.

## 4.2. Algoritmo de segmentación de vasos sanguíneos

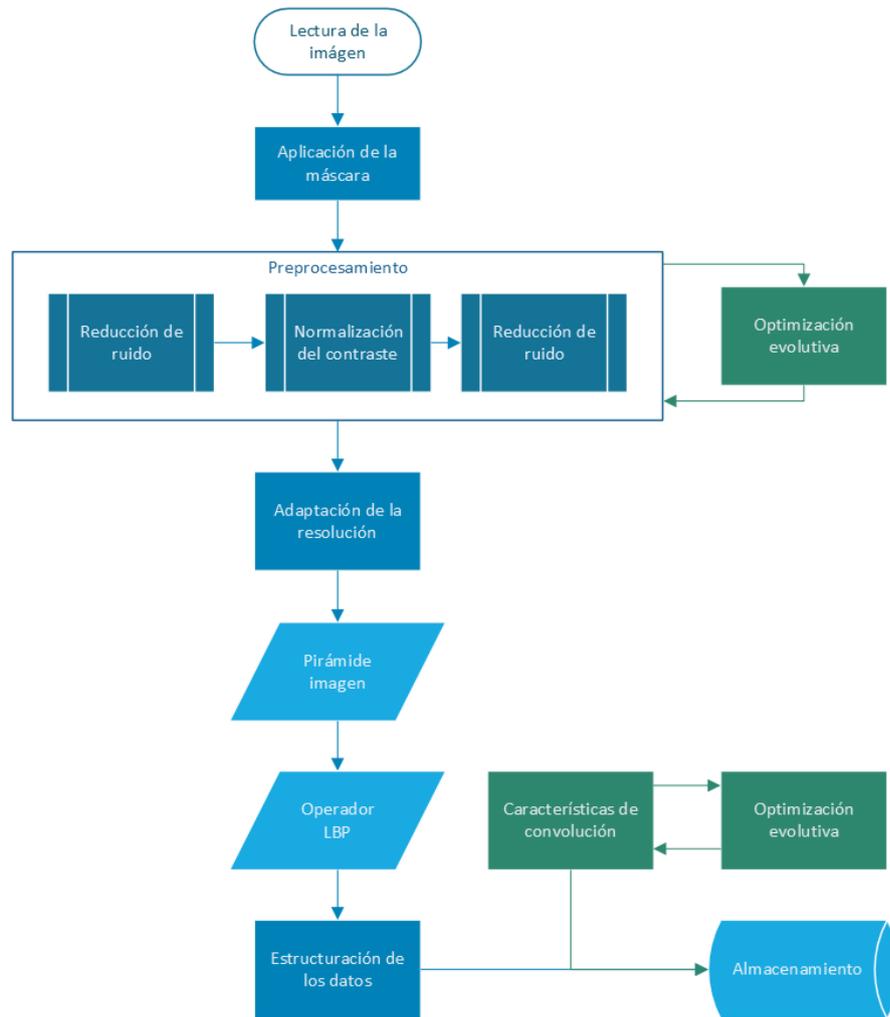
El algoritmo desarrollado con la finalidad de alcanzar los objetivos propuestos en la definición del trabajo (véase la Sec. 1.4) se conforma a partir de dos procesos independientes: la generación del conjunto de datos estructurado a partir de las imágenes de las bases de datos y el entrenamiento de un clasificador a partir de dicho conjunto de datos que sea capaz de realizar la predicción de resultados sobre nuevas imágenes, proporcionando así mismo una representación gráfica de la predicción realizada. En las Fig. 4.4 y 4.5 se proporcionan los diagramas de los procesos indicados, donde se puede diferenciar entre las etapas en color azul, correspondientes al problema de generación de características discriminantes por medio del operador *LBP*, y de color verde, correspondientes al problema de aplicación (véase la Sec. 1.2).

### 4.2.1. Definición de la estructura del dataset

La información utilizada para clasificar los píxeles de las imágenes en el proceso de segmentación parte de imágenes con diferentes resoluciones debido a la pirámide de imagen y, por lo tanto, el número de píxeles existente en cada una de las resoluciones será diferente, dificultando la generación de un conjunto de datos estructurado. En esta sección se detalla la metodología llevada a cabo para su generación.

#### 4.2.1.1. Resolución de las imágenes de partida

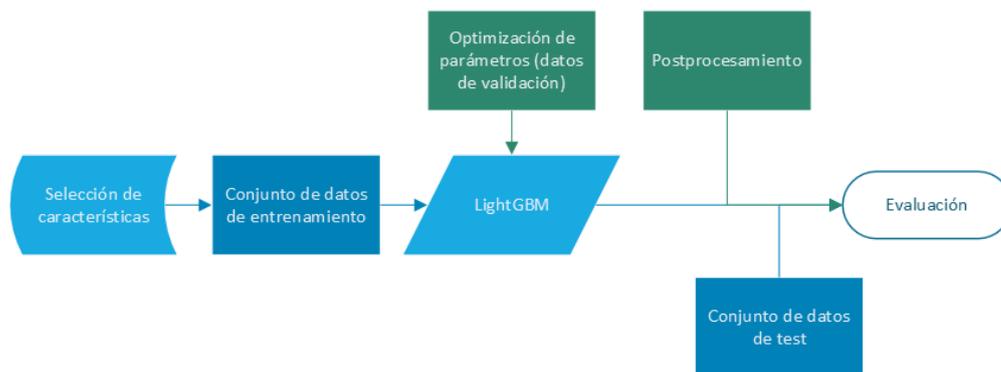
En primer lugar, se debe considerar que en el proceso de reescalado para la generación de la pirámide, en el que se reduce a la mitad el número de píxeles tanto del ancho como del alto de las imágenes en cada iteración, el número final de píxeles de



**Figura 4.4.:** Diagrama del algoritmo de generación del conjunto de datos. En verde se encuentran las etapas de la Sec. 4.2.3.

la imagen de menor resolución debe tener un valor entero. Por lo tanto, la resolución de la imagen original debe ser múltiplo de 2 tantas veces como niveles de pirámide de inferior resolución a la de partida se quieran emplear. Para realizar un análisis lo más detallado posible a lo largo de los experimentos se ha definido en cinco el número de niveles, lo que implica disminuir 1024 veces el número de píxeles de la imagen original. No se ha considerado de interés el análisis de un mayor número de niveles debido a la pérdida de información que supone la reducción de píxeles asociada.

Con un número de niveles igual a cinco, el total de píxeles tanto de ancho como de alto de las imágenes originales debe ser un múltiplo de  $2^5 = 32$ . Para el caso concreto de *DRIVE*, cuya resolución de las imágenes es de  $584 \times 565$ , no se cumple dicha condición, dado que la resolución de las escalas inferiores tendría un valor de



**Figura 4.5.:** Diagrama de entrenamiento y evaluación del clasificador. En verde se encuentran las etapas de la Sec. 4.2.3.

18,25 × 17,66. De entre las diferentes alternativas posibles, la decisión que se ha tomado es la adición de unos «bordes» con píxeles de valor nulo que posteriormente serán filtrados mediante la aplicación de la máscara correspondiente al *campo de visión*. De esta forma se evitan procesos de reescalado sobre la imagen original y se conserva el *groud truth*.

Para realizar esta operación se define la resolución de la imagen del último nivel y con ello se calcula la resolución deseada para la escala 1 : 1, determinando el número de píxeles que deben ser añadidos a la imagen a escala original, a la máscara y al *groud truth*. De esta manera, para alcanzar la resolución escogida se añaden dos «bordes», en la parte superior y en la derecha de la imagen, de píxeles ficticios con valor nulo. Continuando con el ejemplo de dicho dataset,

$$18,25 \times 17,66 \rightarrow 19 \times 18 \rightarrow 19 \cdot 32 \times 18 \cdot 32 = 608 \times 576$$

Para *DRIVE*, por lo tanto, se añade una región superior de dimensión 608 – 584 = 24 → 24 × 576 y una región derecha de dimensión 576 – 565 = 11 → 584 × 11.

El proceso es equivalente *CHASE* y *STARE*, facilitando de esta forma la generación de la pirámide de imagen sin interferir posteriormente al proceso de clasificación o evaluación al conservarse el *groud truth* original, sin ningún reescalado sobre el mismo.

#### 4.2.1.2. Estructura deseada del conjunto de datos

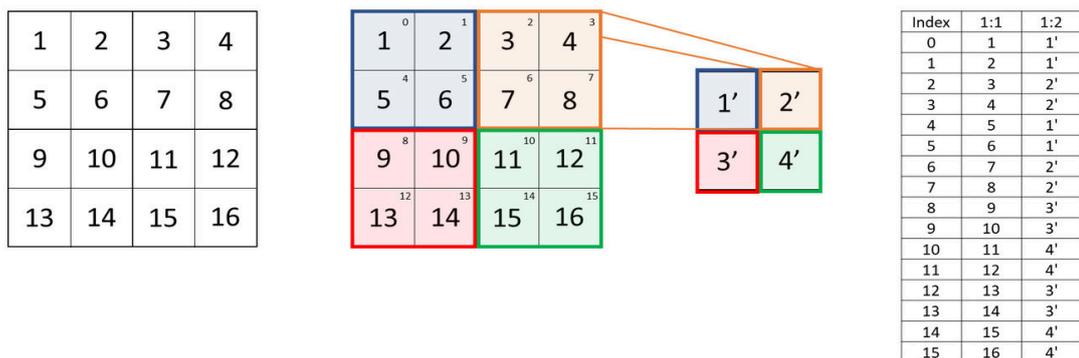
Puesto que el objetivo final del procesamiento y tratamiento de las imágenes es la segmentación por medio de una clasificación a nivel de píxel, se busca una estructura en la que cada píxel representa una fila, es decir, un registro. Se pretende que las columnas sean características de cada píxel provenientes de la información generada por el procesamiento de la imagen original mediante diferentes técnicas y operadores.

Con la estructura mencionada es posible el tratamiento individual de cada uno de los píxeles, descartando aquellos que no formen parte de la máscara tanto en los procesos de entrenamiento como de testeo.

### 4.2.1.3. Mapeo de píxeles entre las diferentes escalas

Como ha sido expuesto en la Sec. 4.2.1.2, las características del conjunto de datos deben describir los píxeles de la imagen original. En la construcción de la pirámide de imagen, proceso previo a la aplicación del operador *LBP*, se ve alterado el número de píxeles y, por lo tanto, se pierde la correspondencia con respecto a la imagen de partida, siendo necesario reestructurar la información de las diferentes resoluciones con respecto a dicha imagen.

De manera intuitiva, se puede asumir que la información de los diferentes niveles de inferior resolución de la pirámide proviene de los píxeles más próximos a la posición del nuevo píxel generado en el proceso de interpolación. Por lo tanto, en el proceso de reescalado, en el que en cada escala el número de píxeles se ve disminuido 4 veces, se asocia la información del nuevo píxel con la de los cuatro píxeles que se encuentran en su posición equivalente en el nivel previo.



**Figura 4.6.:** Generación del conjunto de datos estructurado mediante imágenes a dos escalas diferentes (1 : 1 y 1 : 2).

En la Fig. 4.6 se muestra la referencia descrita en el reescalado de una imagen de 16 píxeles a 4 nuevos píxeles mediante interpolación. Así mismo, se puede observar que por medio de dicha referencia se asignan los valores de los nuevos píxeles a los de la imagen original en el proceso final de generación del conjunto de datos. El valor obtenido en la imagen de inferior resolución se asigna a los cuatro píxeles más próximos de la imagen de mayor resolución, asignando a todos ellos el mismo valor. Este proceso es repetido para escalas de resoluciones inferiores, de modo que, por ejemplo, el valor de cada píxel de la escala 1 : 4 será compartido por 16 píxeles de la escala original.

Puesto que también se planea trabajar con una escala superior a la original, la 2 : 1, se encuentra el problema inverso, un número de píxeles superior a los de la resolución original. El problema se resuelve de forma equivalente, utilizando la referencia entre 4 píxeles de la imagen de mayor resolución con 1 píxel de la imagen de una escala inferior, de forma que cada registro en el dataset continúa siendo un píxel de la escala 1 : 1, y, por lo tanto, en lugar de generar una nueva característica se obtiene un total de 4 adicionales. En la Fig. 4.7 se tiene un ejemplo de construcción del dataset para la escala 2 : 1 a partir de una imagen de 4 píxeles.

Index	1:1	2:1-1	2:1-2	2:1-3	2:1-4
0	1	1'	2'	5'	6'
1	2	3'	4'	7'	8'
2	3	9'	10'	13'	14'
3	4	11'	12'	15'	16'

**Figura 4.7.:** Generación del conjunto de datos para una escala 2 : 1.

Con la metodología descrita se alcanza la estructura deseada para el conjunto de datos que posteriormente será empleado en los procesos de aprendizaje automático y clasificación, es decir, registros que pueden ser tratados de manera independiente entre sí y representativos de píxeles en la escala original de la imagen.

#### 4.2.2. Esquema general de funcionamiento del algoritmo

Los pasos que sigue el algoritmo son los definidos a continuación:

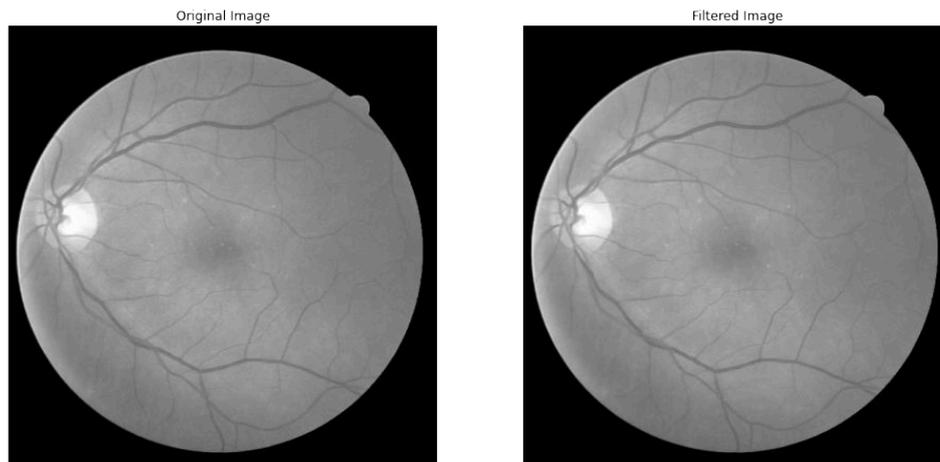
1. Lectura de las imágenes como matrices de valores enteros en función de la intensidad del canal correspondiente o de la combinación de los tres canales.
2. Anulación de los píxeles que no forman parte del fondo de ojo, proporcionándoles un valor nulo, mediante la aplicación de la máscara del *campo de visión*.
3. Debido a las particularidades de cada una de las imágenes de la base de datos *DRIVE*, y a la posibilidad de mejorar las imágenes de partida empleadas en el operador *LBP* a nivel de contraste y de reducción del ruido, se han definido manualmente una serie de procedimientos de preprocesamiento de las imágenes. No obstante, en la Sec. 4.2.3.1 se detalla la metodología usada en posteriores experimentos para eliminar la intervención del usuario a la hora de sintonizar los parámetros de la etapa de preprocesamiento, optimizando automáticamente el valor de los mismos para cada base de imágenes.

Como preprocesamiento de partida se tienen las siguientes fases:

- a) Primera fase de reducción del ruido en la que busca la conservación de los bordes de la imagen, dado que estos delimitan los vasos sanguíneos. Para ello se ha escogido un filtro bilateral que tiene como principal característica la preservación de los bordes, minimizando el efecto sobre los

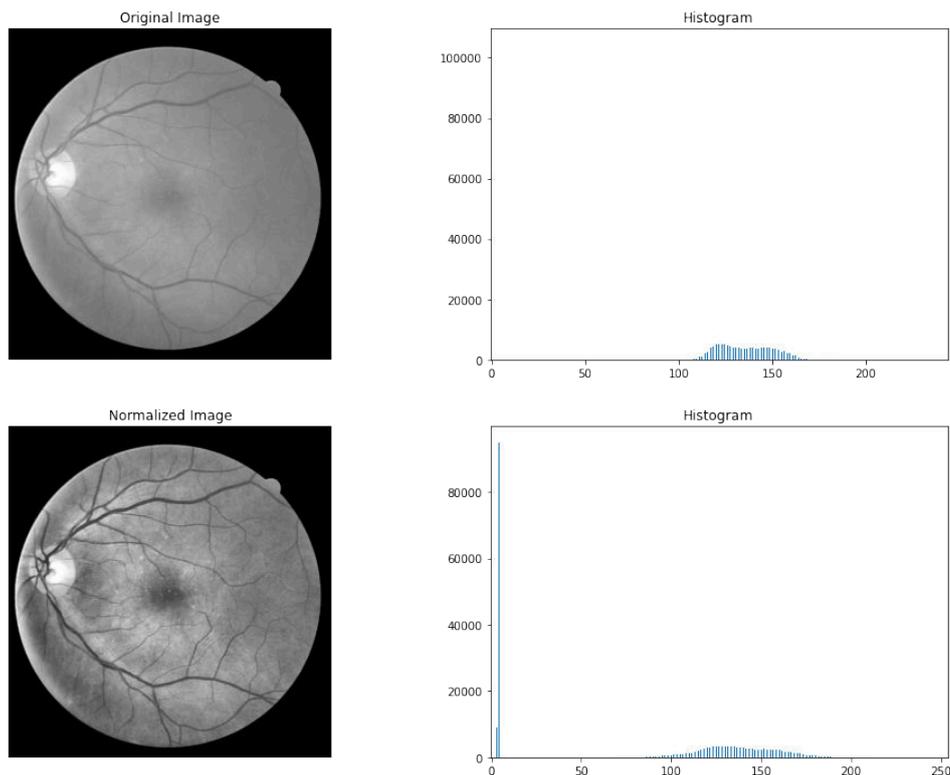
vasos sanguíneos más finos. El filtro bilateral posee tres parámetros [33] que han sido configurados manualmente:

- Diámetro de la vecindad de los píxeles vecinos: el valor escogido es 3.
- Parámetro de color, donde un valor mayor implica que colores más lejanos serán considerados: 3.
- Parámetro relativo al espacio en el que un mayor valor implica que píxeles más lejanos se van a influenciar en mayor medida: 9.



**Figura 4.8.:** Ejemplo de resultado de la primera operación de filtrado para la reducción de ruido en el proceso de preprocesamiento.

- b) Normalización de las imágenes mediante la mejora del contraste local (*CLAHE*). Se realiza mediante el cálculo de los histogramas de intensidades en regiones locales de la imagen, forzando a que se encuentren distribuidos en el rango completo de valores de intensidad disponibles, es decir, entre 0 y 255. De esta forma, se busca resaltar los vasos sanguíneos de manera independiente a la intensidad luminosa de la imagen de partida. En la Fig. 4.9 se muestra un ejemplo de resultado de la aplicación del filtro, así como su efecto genérico sobre el histograma de valores de intensidad en la imagen global, el cual muestra finalmente una mayor uniformidad a lo largo del rango de valores disponible. En dicha figura también es posible visualizar un aumento del contraste que resalta visualmente los vasos sanguíneos. Se han considerado los siguientes parámetros [34]:
- Valor límite para definir las posibles variaciones de contraste: 3.
  - Tamaño de la malla de regiones de la imagen considerada para igualar los histogramas: 8.
- c) Repetición de la reducción de ruido mediante un filtro bilateral.



**Figura 4.9.:** Ejemplo de resultado tras la normalización de la imagen en el proceso de preprocesamiento.

4. Adaptación de la resolución de las imágenes para que sea posible el proceso de reescalado a inferiores resoluciones, generando la pirámide de imagen, donde es preciso que el número de píxeles de la imagen sea múltiplo de 2 tantas veces como «niveles» tenga la pirámide. Para ello se lleva a cabo una adición de «bordes ficticios», introduciendo píxeles con valor de intensidad nulo, para alcanzar la resolución deseada sin llevar a cabo un proceso de interpolación que pierda la correspondencia entre píxeles con el *ground truth*.
5. Reescalado de la imagen a las siguientes escalas:  $2 : 1$ ,  $1 : 2$ ,  $1 : 4$ ,  $1 : 8$ ,  $1 : 16$ ,  $1 : 32$ . El algoritmo empleado para la interpolación correspondiente al reescalado es *Lanczos* [35] debido a la elevada calidad que proporciona tanto en los aumentos de resolución como en las reducciones.
6. Aplicación del operador *LBP* y sus variantes (véase Sec. 3.1.2): *RI*, *RIU2*, *NRI* y *VAR*.
7. Como procedimiento final en lo relativo al tratamiento y aplicación de operadores, los resultados obtenidos son almacenados como variables que contienen matrices de datos con las dimensiones correspondientes a cada escala.
8. Finalmente, para la generación del conjunto de datos que podrá ser directamente utilizado para el entrenamiento, validación o test del clasificador, se



**Figura 4.10.:** Ejemplo de resultado de la segunda operación de filtrado para la reducción de ruido en el proceso de preprocesamiento.

recogen las variables almacenadas en el proceso previo, realizando las operaciones indicadas en el apartado 4.2.1.3 para igualar el número de píxeles entre las diferentes escalas. Nuevamente, se emplea la información de la máscara para descartar aquellos píxeles que no formen parte del fondo de ojo. Dado que la construcción del conjunto de datos estructurado puede ser configurada, se han definido los siguientes parámetros para determinar las características que finalmente lo constituirán:

- Inclusión o no de la escala 2 : 1: permite seleccionar la utilización de dicha escala de modo que un total de cuatro características nuevas sean añadidas a cada registro (véase Fig. 4.7).
- Número de escalas a utilizar: definición del número de niveles de la pirámide de imagen. El parámetro se configura de modo que la escala 1 : 1 siempre es incluida. Así mismo, si el parámetro de inclusión de escala 2 : 1 es verdadero, esta será igualmente incluida. De este modo, el número de escalas inferiores a la escala original quedará definido mediante la operación:

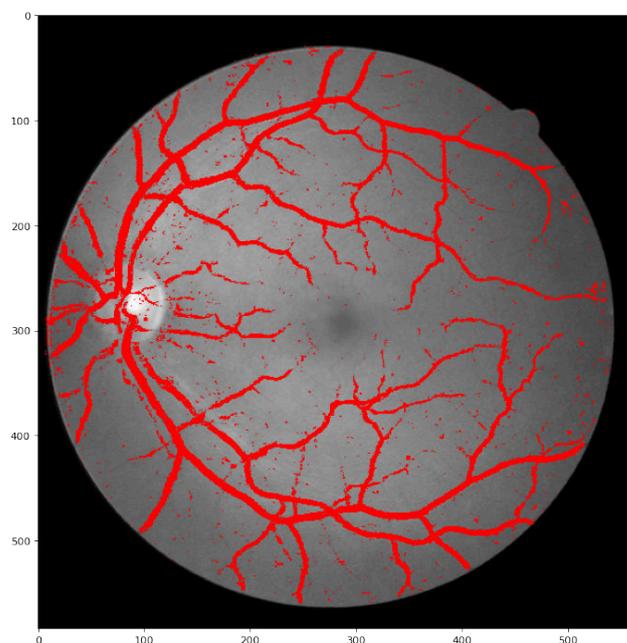
$$n \text{ escalas inferiores} = n \text{ escalas} - (1 + 1 * \text{inclusión escala } 1 : 1)$$

- Inclusión de la intensidad de la imagen a escala original de manera previa a la aplicación de operadores *LBP*.
- Variante del operador *LBP* a utilizar.

Generados los conjuntos de datos que pueden ser directamente empleados en el clasificador, se tiene una segunda fase del algoritmo consistente en el propio proceso

de entrenamiento y evaluación del clasificador, que puede ser configurada mediante las siguientes opciones:

- Variantes del operador *LBP* a utilizar en la constitución final del conjunto de datos.
- Canales de la imagen a emplear.
- Características adicionales. Permite considerar características de procesos independientes con la única condición de que los registros se correspondan con los píxeles de la imagen a escala 1 : 1. Este parámetro es empleado, por ejemplo, en la inclusión de las características generadas en los procesos de convolución descritos en la Sec. 4.2.3.2.
- Visualización de los píxeles etiquetados como vasos. Por medio de este parámetro es posible mostrar los resultados gráficos por medio de la reconstrucción de las imágenes a partir de la máscara.



**Figura 4.11.:** Ejemplo de predicción de píxeles etiquetados como vasos durante la fase de predicción.

### 4.2.2.1. Aplicación del operador *LBP* sobre la pirámide de imagen

Dados los objetivos sobre los que la investigación ha sido definida, una de las hipótesis de partida es la utilización de una pirámide de imagen de manera previa a la aplicación del operador *LBP* con el objetivo de simular la utilización de radios superiores a la unidad en su aplicación, es decir, la equivalencia entre el operador

$LBP$  de radio  $n + 1$ , siendo  $n \geq 0$ , con operador  $LBP$  de radio 1 aplicado al nivel de pirámide  $1/2^n$ .

El primer análisis realizado en el apartado metodológico del trabajo es la definición de la estructura de datos necesaria para realizar la segmentación de imágenes mediante el operador a diferentes escalas y poder realizar así la verificación de la hipótesis de partida. Se emplea por lo tanto la estructura del conjunto de datos establecida en la Sec. 4.2.1.3 para la conformación de los datos estructurados en los que las características, obtenidas por medio del operador  $LBP$ , son referenciadas con respecto a los píxeles de la imagen original, facilitando así el entrenamiento de un clasificador que etiquete individualmente cada uno de los píxeles. Por medio de los resultados de los experimentos de la Sec. 5.3.1 de la aplicación del operador  $LBP$  en diferentes vecindades a través de la pirámide de imagen es posible verificar la hipótesis indicada.

#### 4.2.2.2. Generación de características complementarias

Partiendo así mismo de la información de la imagen original y de la aplicación del operador  $LBP$  con las estructuras de datos definidas en la Sec. 4.2.1.3, se realiza la verificación del incremento del poder predictivo de la utilización de la intensidad de la imagen tras la aplicación del preprocesamiento (véase la Sec. 5.3.3), y de la inclusión de la escala 2 : 1 como nivel de la pirámide de imagen (véase la Sec. 5.3.4).

#### 4.2.3. Otras etapas adicionales

En esta sección se indican procesos adicionales llevados a cabo con la finalidad de mejorar el poder predictivo de los procesos ya definidos y complementarlos con nuevas características. Es decir, definidos los pasos para realizar la experimentación y generar las características correspondientes al problema del trabajo, se proponen metodologías adicionales para el incremento de las prestaciones en el problema de aplicación, la segmentación de vasos sanguíneos.

##### 4.2.3.1. Sintonización de la etapa de preprocesamiento usando computación evolutiva

El preprocesamiento de las imágenes de las bases de datos se constituye mediante las fases definidas en la Sec. 4.2.2, donde se realizan dos fases de reducción de ruido mediante un filtro bilateral junto con una normalización del contraste con el algoritmo  $CLAHE$  que se aplica entre ambos filtros.

Ambos algoritmos, el filtro bilateral y  $CLAHE$ , poseen una serie de parámetros que deben ser ajustados. Concretamente, el filtro bilateral posee un total de 3 parámetros y es empleado en dos ocasiones, y el operador  $CLAHE$  dispone de dos parámetros

adicionales, por lo que se tiene un total de 8 parámetros consistentes en números enteros positivos que deben ser seleccionados. En la Sec. 4.2.2 se indican los parámetros escogidos durante el ajuste manual. El proceso puede verse mejorado mediante su automatización, eliminando un posible error humano y facilitando su aplicación en diferentes bases de datos donde los parámetros deben adaptarse a las propiedades particulares de cada una de ellas.

Dadas las características del problema de optimización y su complejidad, la herramienta de optimización escogida para la selección automática de dichos parámetros es la *Computación Evolutiva* (véase la Sec. 3.4) puesto que permite la optimización mediante un proceso conceptualmente de menor complejidad y de un reducido número de pasos, donde lo más relevante es la definición de una función de evaluación que pueda establecer un criterio de comparación.

Como función de evaluación se emplea un sistema de segmentación de imágenes equivalente al que se está resolviendo en el trabajo, pero simplificado. Concretamente, para evaluar un conjunto de parámetros se toma el canal verde de las imágenes, se le aplica el preprocesamiento definido por los parámetros que se están validando y se realizan los siguientes pasos:

1. Para cada una de las imágenes sobre las que se ha aplicado el preprocesamiento se generan tres características que formarán parte de un conjunto de datos estructurado donde cada píxel se corresponde con un registro:
  - a) El valor de intensidad del píxel.
  - b) El valor *LBP* en su variante *RI* sobre la escala 1 : 2.
  - c) El valor *LBP* en su variante *VAR* sobre la escala 1 : 2.

La razón por la que se aplica el operador *LBP* a escala 1 : 2 es la disminución del tiempo de computación. Al reducir el número de píxeles se reduce el número de transformaciones con el operador. Así mismo, se simula parcialmente el uso de la pirámide de imagen.

2. De las imágenes de entrenamiento de la base de datos, se realiza una subdivisión 70 % - 30 % en conjuntos entrenamiento y validación.
3. A su vez, del subconjunto de entrenamiento se selecciona una muestra del 20 % de píxeles de manera estratificada, es decir, conservando la proporción de píxeles correspondientes a vasos sanguíneos respecto a la proporción de la muestra original. Con ello se busca reducir los tiempos de computación.
4. Aplicación de las máscaras correspondientes al *campo de visión*.
5. Se entrena un clasificador *LightGBM* y se evalúa con la submuestra completa de validación.
6. El valor final de la función de evaluación del algoritmo evolutivo es la métrica *F1* obtenida.

De entre las posibles técnicas de definición del *algoritmo evolutivo*, han sido escogidas las siguientes:

- Individuos compuestos por 8 genes que pueden tomar valores decimales positivos con inicialización aleatoria. Cada uno de ellos se corresponde con uno de los parámetros de preprocesamiento. Dado que los parámetros deben ser empleados como números enteros se realiza un redondeo de los estos como primer paso en la función de evaluación. El tamaño de población escogido es de 50 individuos.
- Selección de padres mediante torneo de 2 individuos. El número de padres es igual al tamaño de la población, de forma que se escogen 50 parejas aleatorias de individuos quedando como padres resultantes los 50 con mejor valor de evaluación de cada una de las parejas.
- Recombinación aritmética con una probabilidad del 60 % de aplicación a parejas de padres emparejadas aleatoriamente. En caso de ser aplicadas se generan dos nuevos individuos mediante la siguiente operación:

$$descendiente_1 = padre_1 + (1 - \alpha) * padre_2$$

$$descendiente_2 = padre_2 + (1 - \alpha) * padre_1$$

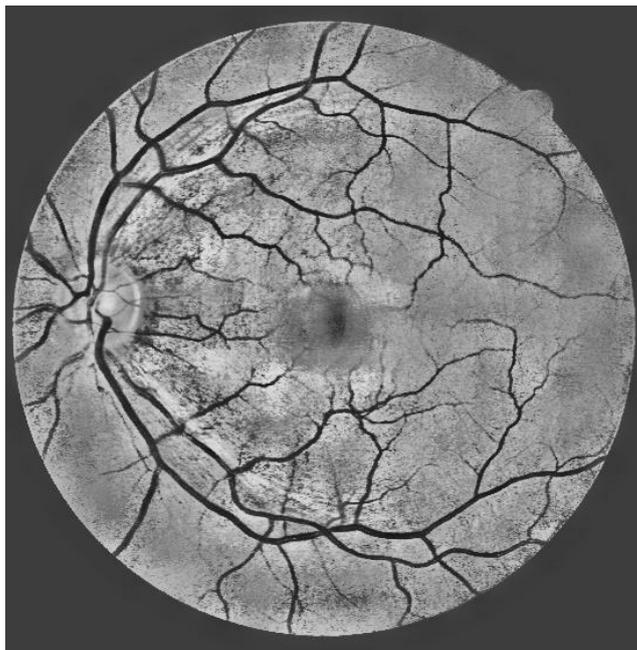
Se toma un parámetro  $\alpha = 0,6$ . Los descendientes serán exactamente iguales que los padres si debido a la probabilidad establecida no se aplica la recombinación.

- Mutación no correlacionada basada en una distribución normal de media nula y  $\sigma = 0,4$ . La mutación se aplica con una probabilidad del 15 % a nivel de gen.
- Elitismo en la selección de la descendencia final que constituirá la población de la siguiente iteración. Al aplicar elitismo, el individuo de la descendencia con peor valor de evaluación es reemplazo por el mejor individuo de la nueva generación. Esta condición no aplica cuando el valor de evaluación de todos los individuos de la nueva generación es superior al de todos los individuos de la generación previa.

La principal razón para la utilización de una representación de los individuos de valores reales es la compatibilización y facilitación de la reutilización del *algoritmo evolutivo* diseñado para el proceso de generación de características mediante convolución descrito en la próxima sección (véase la Sec. 4.2.3.2).

#### 4.2.3.2. Generación de características mediante convolución

Se propone la inclusión de imágenes convolucionadas como etapa adicional en la que obtener una serie de características complementarias que permitan incrementar el poder predictivo. De esta forma se busca aproximar el proceso en la medida de lo



**Figura 4.12.:** Resultado del preprocesamiento sobre el canal verde en una de las imágenes del dataset *DRIVE*.

posible a resultados que resulten competitivos con las investigaciones más recientes de la literatura en las que se emplean las mismas bases de imágenes mediante metodologías no basadas en *Deep Learning*.

Se trata de un acercamiento a la metodología llevada a cabo por las redes neuronales convolucionales en las que se aplican convoluciones optimizadas automáticamente, pero sin llegar a la utilización de *Deep Learning* y las desventajas descritas en la Sec. 1.4. Con ello, y en línea con las metodologías descritas en fases previas del trabajo, se lleva a cabo la inclusión de seis características adicionales consistentes en el valor de intensidad de los píxeles en seis imágenes generadas mediante la aplicación de seis núcleos de convolución sobre el canal verde a cada una de las retinografías en escala original.

Con la finalidad de que las imágenes resultantes de las convoluciones puedan aportar información diferente pero complementaria entre sí, se ha decidido emplear un total de seis núcleos de diferentes dimensiones, siendo dos de ellos de tamaño  $3 \times 3$ , dos de tamaño  $5 \times 5$  y otros dos adicionales de dimensión  $7 \times 7$ . El proceso de optimización se lleva a cabo de forma simultánea de modo que la información conjunta de las 6 características proporcione el mejor resultado posible, siendo esperable de esta forma, por lo tanto, que la información que aporta cada una de las seis imágenes resultantes no sea equivalente a la de las cinco restantes y que de manera conjunta proporcionen la mayor cantidad de información posible. El criterio de utilización de 3 dimensiones diferentes es el facilitar el análisis de vecindades de diferentes tamaños.

Los núcleos de convolución se obtienen mediante un proceso de optimización auto-

matizado por medio de un *algoritmo evolutivo* como el descrito en la Sec. 4.2.3.1, en la que se indica la metodología de generación automática del preprocesamiento de las imágenes. El *algoritmo evolutivo* consta de las mismas etapas, con las siguientes particularidades:

- Utilización de individuos de 166 genes compuestos por números reales que pueden ser tanto positivos como negativos, donde cada uno de ellos es uno de los elementos de las 6 matrices de convolución. Inicialización aleatoria de valores en el intervalo  $[-1, 1]$ .
- Sistema equivalente de selección de padres, recombinación, mutación y elitismo. Los parámetros escogidos que difieren con los indicados en el apartado previo son un tamaño de población de 100 individuos y  $\sigma = 0,1$  para la mutación.

En lo relativo a la función de evaluación, se realizan los pasos:

1. Lectura del canal verde de las imágenes del conjunto definido como entrenamiento para cada una de las bases de datos de retinografías.
2. Para el individuo dado, el cual define un total de seis núcleos de convolución, aplicación de cada una de las correspondientes convoluciones a cada una de las imágenes.
3. División de las imágenes convolucionadas a su vez en un 70 % de entrenamiento y un 30 % de validación.
4. Del subconjunto de entrenamiento se selecciona una muestra del 20 % de píxeles de manera estratificada.
5. Aplicación de las máscaras (*campo de visión*).
6. Se entrena un clasificador *LightGBM* con la finalidad de identificar los píxeles de vasos sanguíneos y se realiza una predicción y evaluación sobre la submuestra de validación.
7. El valor final de la función de evaluación es la métrica *F1* obtenida en el proceso.

#### 4.2.3.3. Optimización de parámetros del clasificador

El clasificador empleado para el entrenamiento y la predicción de los píxeles pertenecientes a un vaso sanguíneo posee una serie de parámetros que pueden ser manualmente configurados u optimizados mediante una búsqueda exhaustiva realizada sobre valores discretos de los parámetros del modelo para que el entrenamiento se vea correctamente adaptado a los datos y de dicha forma incrementar el poder predictivo del mismo.

Inicialmente, se han empleado una serie de parámetros definidos mediante un proceso de selección manual basado en las propiedades del problema y a iteraciones

manuales de sintonización mediante la comprobación de los resultados obtenidos. Los parámetros manualmente configurados son los siguientes:

- Número máximo de hojas, es decir nodos terminales, para cada uno de los árboles. El parámetro por defecto es 31 mientras que el valor empleado es 50. Mediante un mayor número de nodos terminales es posible lograr un mayor ajuste en el proceso de entrenamiento y adaptarse en mayor medida a la complejidad de los datos, pero incrementando el riesgo de sobreajuste. En el dataset *DRIVE* existe un elevado grado de similitud entre los datos de entrenamiento y evaluación (las imágenes tienen propiedades similares) y por lo tanto el riesgo de sobreajuste es inferior.
- Máxima profundidad, en cuanto a número de nodos, de cada uno de los árboles. El parámetro por defecto es  $-1$  que indica que no existe un límite definido, siendo 30 el valor empleado. El valor ha sido definido principalmente mediante un proceso iterativo de verificación manual.
- Muestreo de las columnas por árbol. Se ha empleado un valor del 90 %, es decir, en cada árbol se descartan aleatoriamente el 10 % de las características disponibles.
- Muestreo del número de registros para los diferentes árboles. Se ha establecido en un 70 % dado que el número de píxeles disponibles es elevado.
- Ratio de aprendizaje de valor 0,5. El valor por defecto es 0,1 y ha sido ajustado mediante iteraciones manuales.

A pesar de que se ha realizado una selección manual de los parámetros, se define un proceso búsqueda exhaustiva con la finalidad de comprobar los resultados de manera conjunta para los diferentes valores de los parámetros, puesto que no son independientes entre sí, logrando así mismo la adaptación automática a las diferentes bases de datos disponibles. El proceso de optimización y selección de los parámetros escogido es el conocido como *Grid Search*, consistente en la definición de un listado de posibles valores para cada uno de los parámetros a optimizar y la verificación de los resultados proporcionados por cada una de las posibles combinaciones de valores. La comparación se realiza mediante los resultados de evaluación en conjuntos de datos de validación de la métrica *F1*.

El conjunto de parámetros definido es el que se indica a continuación:

- Número máximo de hojas: 40, 50 y 60.
- Máxima profundidad: 20, 30 y 40.
- Muestreo de las columnas por árbol: 70 % y 90 %.
- Muestreo del número de registros: 50 % y 70 %.
- Ratio de aprendizaje: 0,2, 0,5 y 0,7.

#### 4.2.3.4. Postprocesamiento automático mediante la información de la vecindad

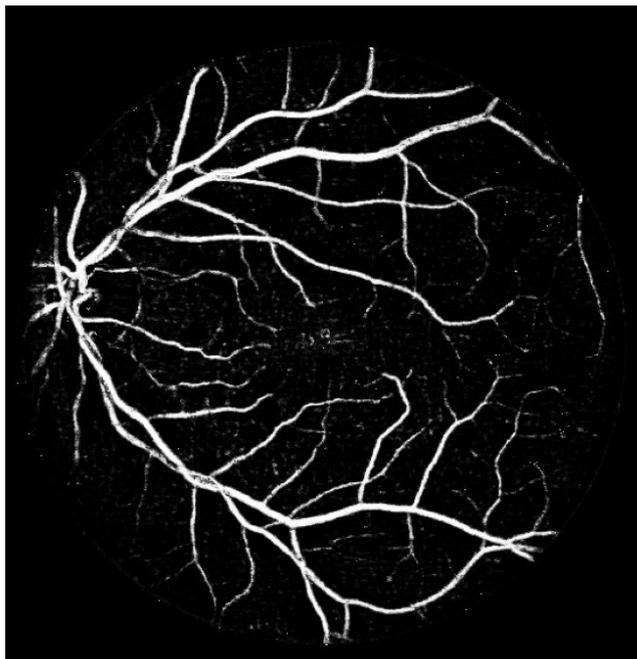
El proceso de segmentación se lleva a cabo mediante una clasificación individual de los píxeles, que, aunque son tratados de manera individual e independiente realmente forman parte de imágenes, lo que implica que los diferentes registros existentes en el conjunto de datos estructurados no son independientes entre sí, sino que los objetos a segmentar, concretamente los vasos sanguíneos, poseen una continuidad que se traduce en una relación entre píxeles que no está siendo considerada. Por ejemplo, si un píxel se encuentra clasificado como vaso sanguíneo mientras que a su vez está rodeado en su vecindad de un elevado número de píxeles cuya predicción es la clasificación opuesta, es altamente probable que la predicción realizada sobre dicho píxel haya sido etiquetada erróneamente. Es decir, una vez finalizado el proceso de predicción sobre cada uno de los píxeles de manera individual e independiente, es posible llevar a cabo una segunda iteración en la que se consideren los resultados de las vecindades de cada píxel para generar un nuevo resultado de clasificación en el que el ruido se vea disminuido y se observe una mayor continuidad en los objetos segmentados.

Dado que de manera previa al entrenamiento y predicción por medio del clasificador se realiza la aplicación de la máscara que define el *campo de visión* de las retinografías, es decir, aquellas regiones de la imagen que forman parte propiamente del fondo de ojo en sí en el proceso de captura de las retinografías, es preciso realizar el proceso inverso, la reconstrucción de la imagen a partir de las predicciones realizadas sobre los píxeles, de forma equivalente a la empleada en la generación los resultados gráficos, es decir, utilizando de nuevo las máscaras para identificar aquellos píxeles que fueron descartados.

A su vez, el grado de fiabilidad de la clasificación de cada píxel como vaso sanguíneo puede ser medido a través de la probabilidad que proporciona el clasificador en lugar de la clasificación binaria inicialmente empleada. En la Fig. 4.13 se encuentra representada como imagen la probabilidad asociada por el clasificador por medio de las características proporcionadas para cada píxel de manera individual, donde una mayor intensidad en escala de grises, es decir, un tono más claro, se asocia con una probabilidad más elevada.

Con la finalidad de incluir información de la vecindad de cada uno de los píxeles acerca de la probabilidad estimada de formar parte de un vaso sanguíneo, se genera para cada uno de ellos una serie de nuevas características que proporcionan información por medio de operaciones de convolución, de reducción de ruido y del propio operador *LBP*. De esta forma, en la segunda iteración de ajuste del clasificador se parte del mismo conjunto de variables que las empleadas en el previamente ajustado, junto con las listadas a continuación:

- Convolución con la que se pretende recoger información detallada de los píxeles inmediatamente próximos por medio de una media aritmética de sus



**Figura 4.13.:** Probabilidad asociada a cada píxel de formar parte de un vaso sanguíneo. Los tonos más blancos indican una mayor probabilidad.

valores. Indica un «grado de fiabilidad» de la región. Por lo tanto, el núcleo de convolución empleado es:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

- Resultado de la aplicación de un filtro gaussiano de dimensión  $9 \times 9$ . El resultado esperado es captar información de una vecindad de mayor tamaño, proporcionando una relevancia superior a los píxeles más próximos, es decir, una operación equivalente a la reducción de ruido de una imagen.
- Transformación de la imagen mediante el operador *LBP* en su versión *VAR*, variante a la intensidad, a escala original de la imagen. El objetivo es recoger información acerca de la varianza de los valores de probabilidad de cada uno de los píxeles en su vecindad más cercana.
- Transformación de la imagen mediante el operador *LBP* en su variante *VAR* a escala  $1 : 2$  de la imagen. Se lleva a cabo con la misma finalidad que a escala  $1 : 1$ , pero con una vecindad de mayor dimensión.
- Aplicación del operador *LBP* en su variante *VAR* a escala  $1 : 4$ .
- Aplicación del operador *LBP* en su versión *RIU2*, invariante a la intensidad, pero variante a la rotación, a escala original. El objetivo es recoger información

así mismo de la geometría de los valores de probabilidad de cada uno de los píxeles en su vecindad más cercana.

- Transformación de la imagen mediante el operador  $LBP$  en su variante  $RIU2$  a escala 1 : 2. Se emplea con la misma finalidad que a escala original, pero con una vecindad más grande.
- Aplicación del operador  $LBP$  en su variante  $RIU2$  a escala 1 : 4 de la imagen.

## 5. Experimentos, resultados y discusión

En este capítulo se proporciona la experimentación, sus correspondientes resultados intermedios y los finales de la investigación, y una discusión sobre los mismos. En el proceso de experimentación y constitución de la metodología se ha empleado únicamente un dataset, *DRIVE*, mientras que en la evaluación en procesos de segmentación de retinografías se han utilizado tanto dicho dataset como adicionalmente las bases de datos *STARE* y *CHASE*. Los resultados obtenidos en *DRIVE* son verificados mediante un conjunto de imágenes de evaluación que ha sido mantenido de manera completamente independiente a los restantes procesos de experimentación. La razón por la que se ha complementado el proceso de evaluación con nuevas bases de imágenes de retinografías es la verificación de la posibilidad de utilización de la metodología en imágenes con características similares, pero no iguales. Así mismo, con ello se facilita el análisis de la competitividad con respecto a la literatura existente. De esta forma, se utilizan imágenes independientes al proceso de entrenamiento como conjunto de evaluación, mediante retinografías no observadas previamente, tal y como ocurriría en aplicaciones reales.

En la literatura analizada en el Cap. 2, correspondiente al *Estado de la Cuestión*, la metodología de evaluación no es siempre correctamente detallada ni dicha independencia alcanzada. Concretamente, en la investigación [16] no se mantiene la independencia, ya que se emplea el 1% de los datos de test para el entrenamiento en la base de datos *STARE*. Así mismo, otras investigaciones como [14, 16, 18] realizan un proceso de evaluación diferente al tomado como «estándar» en la mayoría de las investigaciones, por lo que la comparación no sería realizada exactamente bajo las mismas condiciones que en este trabajo ni la comparación con ellas, por lo tanto, equitativa.

Por último, una ventaja de la utilización de una selección de imágenes como conjunto de evaluación es que facilita el análisis visual de los resultados mediante la representación gráfica de las etiquetas de los píxeles sobre las imágenes originales.

## 5.1. Consideraciones sobre los conjuntos de entrenamiento, de validación y de test

En cualquier proceso de aprendizaje automático supervisado, las bases de datos deben ser divididas en conjuntos de información de entrenamiento, que generalmente se ven sucesivamente divididos en subconjuntos de datos de entrenamiento y validación, y de test. En lo relativo a este trabajo, dado que las propiedades de las bases de datos son diferentes, la metodología de división de cada una de ellas también lo es, siguiendo en la medida de lo posible los estándares definidos en los procesos de investigación previamente realizados.

En lo relativo a *DRIVE*, se emplea la distribución propuesta en la documentación de la base de datos, es decir, la utilización de 20 imágenes para el entrenamiento y 20 para el proceso de test. Dado que se trata de la base de datos empleada en la definición de la metodología de la investigación realizada y donde, por lo tanto, se ha llevado a cabo la experimentación y la verificación de las diferentes hipótesis, se ha realizado una división en el conjunto de imágenes de entrenamiento con una distribución del 70 % para los procesos de entrenamiento y sintonización y el 30 % para la validación, escogiendo las 14 primeras imágenes como conjunto propiamente de entrenamiento y las 6 restantes como validación. Los resultados finales son proporcionados con un entrenamiento sobre el total de 20 imágenes y una evaluación sobre las 20 de test que han sido mantenidas de manera completamente independiente al resto de procesos. Como última consideración para *DRIVE*, se tiene la utilización del *ground truth* correspondiente al primer observador y la utilización de la máscara para el *campo de visión* proporcionada por la base de datos.

Respecto a las bases de datos de *STARE* y *CHASE*, empleadas sobre una metodología previamente definida, que será la obtenida mediante la experimentación, y que, por lo tanto, no incluye dicho proceso iterativo de análisis sobre un conjunto de validación, la división entrenamiento-validación se realiza únicamente para la ejecución de los procesos de optimización. En lo relativo a *CHASE* se emplean las 20 primeras imágenes a modo de entrenamiento, divididas a su vez en 14 y 6 para procesos que precisen de una validación, y las 8 últimas para el test sobre el que se proporcionan los resultados finales. El caso de la base de datos *STARE* resulta más complejo y se emplea la metodología observada generalmente en las investigaciones previas, consistente en la técnica de validación cruzada conocida como *leave-one-out* en la que cada uno de los *folds* corresponde a una imagen, realizando los procesos de ajuste y entrenamiento sobre las 19 imágenes restantes (con división en 13 y 6 imágenes para los procesos de validación de las optimizaciones) en un total de 20 procesos independientes entre sí, proporcionando como resultado final la media de los mismos. La particularidad por la que se emplea dicho proceso en esta base de datos es el reducido número de imágenes. En ambas bases de datos se ha empleado el primer observador como *ground truth* y se ha definido un proceso de generación de las máscaras que determinan el campo de visión:

- Para *CHASE*, el proceso de definición del campo de visión consiste en la aplicación de un filtro en el que los píxeles con intensidad superior a un valor de 5, en una escala donde el valor máximo es 255, son sustituidos por valor 255, y una operación morfológica de apertura con un núcleo de valores unitarios y dimensión  $10 \times 10$ . Dado que los «bordes» de la retinografía que no corresponden al fondo del ojo son completamente negros, el primer filtro permite seleccionar píxeles que con seguridad forman parte del fondo del ojo y proporcionarles un valor que los diferencie plenamente de aquellos que deban ser descartados por el campo de visión. Así, se obtiene una aproximación prácticamente final para el objetivo deseado, la operación de apertura se realiza para rellenar de valor blanco aquellos píxeles más oscuros que hayan podido quedar excluidos.
- En las imágenes de *STARE* se proporciona una metodología alternativa a la mencionada para el anterior conjunto dado que sus propiedades son diferentes y se observan mayores diferencias entre imágenes, como, por ejemplo, en el contraste. En primer lugar, se aplica un núcleo de convolución de valores constantes y dimensión  $5 \times 5$  con la finalidad de difuminar y homogeneizar el fondo del ojo. Posteriormente se aplica un filtro de *CLAHE* que permite la normalización del contraste. A continuación, se realiza un incremento del contraste y aumento del brillo para que los píxeles correspondientes a la retina queden con la mayor intensidad posible. Dada la disparidad de los resultados entre las diferentes imágenes, y conocido que aproximadamente un 25 % de los píxeles no corresponden al campo de visión, se emplea el percentil 26 para definir el valor límite a partir del cual se aplica un filtro por el que los píxeles quedan anulados en cada una de las imágenes.

## 5.2. Métricas de evaluación de los resultados

La métrica escogida para las verificaciones de las hipótesis, las validaciones de los experimentos y, en general, como referencia para establecer comparaciones, y que, por lo tanto, cobra mayor relevancia en este trabajo es la métrica *F1-score*. La razón de su elección se basa principalmente en el desbalanceo existente en las dos posibles etiquetas existentes para cada píxel en el aprendizaje supervisado, donde la amplia mayoría de los píxeles no se corresponden con vasos sanguíneos. La utilización de métricas como el porcentaje de píxeles correctamente etiquetados puede conllevar a resultados equívocos dado que, por ejemplo, la clasificación de todos los píxeles como no pertenecientes a vasos sanguíneos permite alcanzar valores del 80 % – 90 % de precisión cuando en realidad no se están segmentando las imágenes.

Mediante la métrica *F1-score* es posible representar la relevancia correspondiente a la identificación de los vasos sanguíneos para un correcto proceso de evaluación de los resultados dado que se trata de una medición ponderada entre la tasa de predicciones positivas correctamente realizadas y la tasa de píxeles realmente po-

sitivos correctamente clasificados, contando con un equilibrio que métricas como especificidad y sensibilidad no son capaces de proporcionar de manera individual.

De manera adicional, para establecer una comparación con los resultados obtenidos en las restantes investigaciones analizadas en la literatura, se emplean métricas adicionales como la matriz de confusión, la tasa de registros correctamente clasificados (precisión), la especificidad y la sensibilidad. De manera formal, los índices mencionados se definen como se indica a continuación:

$$F1 = \frac{VP}{VP + \frac{1}{2}(FP + FN)}$$

$$Precisión = \frac{VP + VN}{VP + VN + FP + FN}$$

$$Especificidad = \frac{VN}{VN + FP}$$

$$Sensibilidad = \frac{VP}{VP + FN}$$

donde  $VP$ , son los verdaderos positivos;  $VN$ , los verdaderos negativos;  $FP$ , los falsos positivos; y  $FN$ , los falsos negativos

Como consideración final, a lo largo de la sección de experimentación (véase la Sec. 5.3) los resultados gráficos se muestran sobre una de las imágenes de validación de *DRIVE*, representando en color rojo los píxeles correspondientes a las predicciones de vasos sanguíneos sobre la imagen original, pero, sin embargo, las métricas se obtienen mediante el cálculo sobre el conjunto global de píxeles de todas las imágenes de validación. En la Sec. 5.4 se proporcionan los resultados correspondientes a todas las imágenes de test, tanto de manera gráfica como en métricas.

### 5.3. Experimentos realizados

En esta sección se detallan los experimentos realizados en el proceso de definición de la metodología del trabajo.

### 5.3.1. Pirámide de imagen con el operador *LBP*

En esta primera fase de análisis, en la que se desea verificar las bondades de la utilización de la pirámide de imagen, se proporcionan los resultados del entrenamiento y evaluación de clasificadores en los que el conjunto de datos está formado por el resultado del operador *LBP* en su variante original aplicado sobre las imágenes pre-procesadas y reducidas a diferentes escalas. Se parte de un conjunto de datos con una única característica, la correspondiente a la escala 1:1, para añadir sucesivamente características adicionales donde cada una representa una escala, es decir, un nivel de la pirámide de imagen.

De esta forma, cada uno de los clasificadores es entrenado y evaluado con la información de los clasificadores previos y la inclusión adicional de una característica correspondiente a una nueva escala, hasta un total de cinco escalas inferiores, comprobando así si la adición de nuevas escalas aporta mayor poder predictivo.

En el Apéndice A.1 se muestran los resultados gráficos en una de las imágenes de validación tras el entrenamiento de los diferentes clasificadores con la metodología descrita. En particular, la subimagen de arriba a la izquierda de la Fig. A.3 representa el clasificador entrenado con la escala 1 : 1, la imagen de su derecha el entrenado con las dos escalas 1 : 1 y 1 : 2, y así sucesivamente hasta incluir un total de cinco escalas diferentes además de la original. Los resultados gráficos de predicción se muestran únicamente sobre la primera de las 6 imágenes que componen el conjunto de datos de validación, pero, sin embargo, las métricas son calculadas sobre el conjunto global de píxeles de dichas imágenes. En la figura se representan en color rojo los píxeles clasificados como vasos sanguíneos, donde se observa que la inclusión de múltiples escalas, y por lo tanto, de nuevas características, permite aumentar el poder predictivo en el proceso de segmentación de la imagen.

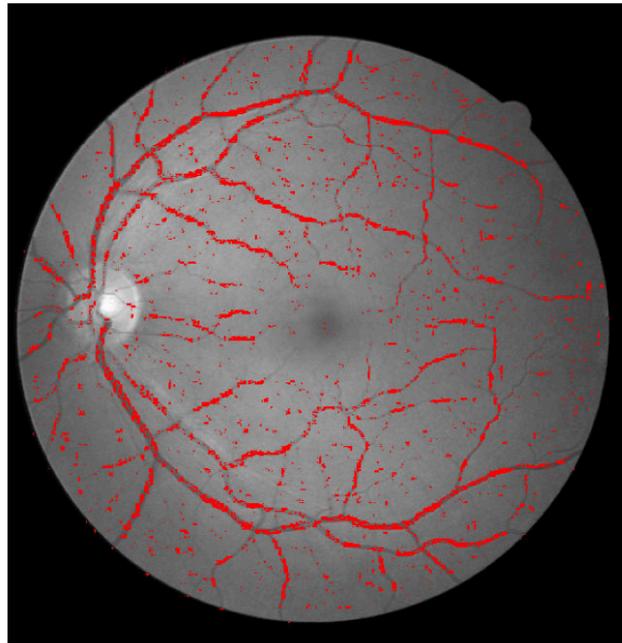
En la Fig. 5.1 se tiene una representación de la aplicación de uno de los clasificadores entrenados, concretamente el correspondiente al empleo de un total de cuatro escalas, tras la aplicación del operador *LBP* en su variante por defecto sobre una de las imágenes del conjunto de validación del dataset *DRIVE*.

Núm. total de escalas	1	2	3	4	5	6
<b>F1-score</b>	0	0,108	0,354	0,420	0,415	0,402
<b>Precisión</b>	0,871	0,873	0,884	0,891	0,889	0,887
<b>Sensibilidad</b>	0	0,060	0,246	0,307	0,306	0,296
<b>Especificidad</b>	1	0,993	0,979	0,977	0,975	0,974

**Tabla 5.1.:** Métricas de evaluación de los resultados de la pirámide de imagen con la aplicación del operador *LBP*. Como ejemplo, *Núm. total de escalas* = 4 indica la utilización simultánea de las características de las escalas 1 : 1, 1 : 2, 1 : 4 y 1 : 8.

De la misma forma, en la Tabla 5.1 se incluyen los resultados de las métricas de

evaluación para los correspondientes clasificadores, extrayendo la misma conclusión que la obtenida mediante el resultado gráfico, el incremento del poder predictivo con la pirámide de imagen. Sin embargo, se puede observar de manera adicional cómo la inclusión de la quinta escala, es decir, la escala 1 : 16, resulta de poca relevancia, posiblemente por la inclusión de ruido generado por la aplicación de un reescalado que modifica considerablemente la imagen original en el proceso de interpolación, incluso perjudicando el poder predictivo del clasificador, efecto que se ve acentuado con la última de las escalas. Por lo tanto, se considera de interés la selección de características que eviten la generación de ruido que reduzca las prestaciones del clasificador. Dicho proceso se llevará a cabo en la Sec. 4.2.2.2.



**Figura 5.1.:** Representación gráfica de la predicción sobre una imagen de validación empleando las características generadas por 4 escalas. Los píxeles rojos han sido etiquetados como vasos sanguíneos.

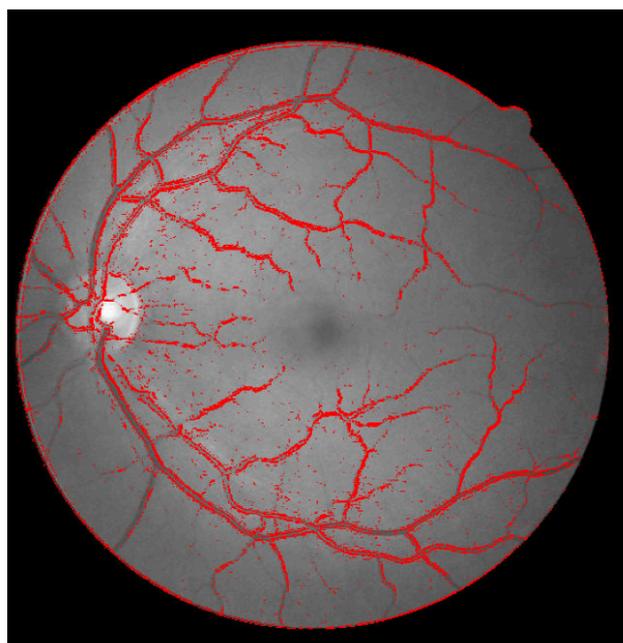
### 5.3.2. Pirámide de imagen con diferentes variantes del operador LBP

Dadas las diferentes variantes del operador *LBP* existentes, las cuales se encuentran analizadas en la Sec. 3.1.2 e incluyen versiones, por ejemplo, invariantes a la rotación o variantes con la intensidad de los píxeles circundantes, es posible extraer diferentes propiedades de los píxeles y completar la información obtenida, de forma que se mejore el poder predictivo del clasificador generado con respecto al de la utilización del operador *LBP* por defecto.

En este apartado se incluye un análisis de los resultados obtenidos de la combinación de la totalidad de variantes del operador *LBP* a múltiples escalas. En la Tabla 5.2 se encuentran dichos resultados, donde la primera columna corresponde con la generación de características por medio de todos los operadores *LBP* a la escala original de la imagen, la segunda incluye dichas características junto con las correspondientes a la escala 1 : 2, y así sucesivamente hasta alcanzar un total de 6 escalas combinadas. Se observa un incremento relevante de las prestaciones de los clasificadores generados con respecto a la utilización del operador original de manera exclusiva o con múltiples escalas de manera simultánea (véase la Tabla 5.1).

Núm. total de escalas	1	2	3	4	5	6
<b>F1-score</b>	0,393	0,561	0,634	0,652	0,650	0,654
<b>Precisión</b>	0,878	0,901	0,914	0,916	0,915	0,916
<b>Sensibilidad</b>	0,306	0,495	0,578	0,613	0,616	0,622
<b>Especificidad</b>	0,963	0,960	0,964	0,961	0,959	0,959

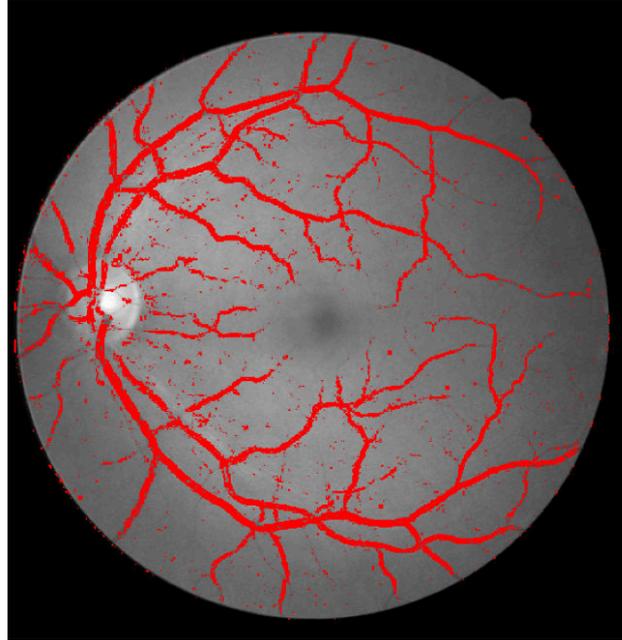
**Tabla 5.2.:** Métricas de evaluación de los resultados de la aplicación de las diferentes variantes del operador *LBP* a múltiples escalas.



**Figura 5.2.:** Resultado gráfico de la aplicación de las diferentes variantes del operador *LBP* a escala 1 : 1.

En la Fig. A.4 se puede observar una segmentación relevante de los vasos sanguíneos, quedando patentes las ventajas proporcionadas por la utilización de múltiples escalas de manera simultánea, donde en la utilización únicamente de la escala 1 : 1, es decir, en la subimagen superior izquierda (representada en mayor dimensión en la Fig. 5.2),

puede ser observado el problema del uso exclusivo de vecindades de tan pequeño diámetro, dado que la región central de los vasos sanguíneos de mayor dimensión no se encuentra correctamente etiquetada, únicamente los bordes.



**Figura 5.3.:** Representación gráfica de la predicción sobre una imagen de validación empleando las características generadas por 4 escalas con múltiples variantes del operador *LBP* simultáneamente.

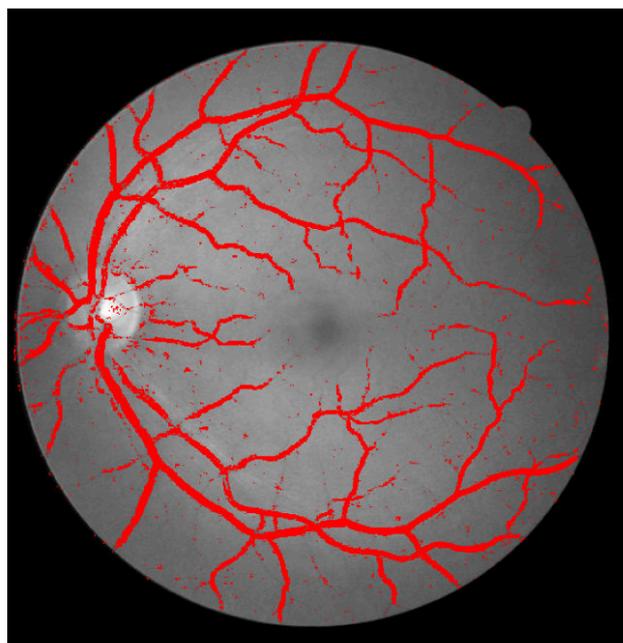
### 5.3.3. Introducción de la intensidad del nivel de gris como característica

Las imágenes originales, es decir, sin la aplicación de operadores y a escala 1 : 1, poseen información que es de utilidad para llevar a cabo el proceso de clasificación dado que, por ejemplo, se trata de la información empleada por un experto para realizar un análisis manual sobre el que generar la segmentación. Por lo tanto, se incluye una característica adicional que represente dicha información. Sin embargo, con el fin de homogeneizarla entre las diferentes imágenes de las bases de datos, se incluye con el preprocesamiento aplicado.

En la Tabla 5.3 y en las Fig. 5.4 y A.5 se encuentran los resultados correspondientes al proceso descrito. La inclusión de la característica adicional proporciona resultados con mejor rendimiento a los obtenidos en la fase previa en la que dicha información no ha sido incluida (véase la Tabla 5.2).

Núm. total de escalas	1	2	3	4	5	6
<b>F1-score</b>	0,570	0,661	0,690	0,699	0,692	0,687
<b>Precisión</b>	0,898	0,917	0,924	0,925	0,923	0,921
<b>Sensibilidad</b>	0,523	0,628	0,659	0,680	0,676	0,677
<b>Especificidad</b>	0,954	0,960	0,963	0,961	0,959	0,956

**Tabla 5.3.:** Métricas de evaluación tras la inclusión de la característica de intensidad de las imágenes con el preprocesamiento aplicado.



**Figura 5.4.:** Representación gráfica de la predicción sobre una imagen de validación con la adición de la intensidad.

#### 5.3.4. Utilización de la escala 2:1

En esta sección se evalúa la inclusión de características generadas mediante la aplicación del operador *LBP* sobre una escala de dimensión superior a la original de la imagen. Con la generación de imágenes a inferior resolución se facilita el análisis de vecindades de mayor dimensión que la de un radio unitario, sin embargo, con la utilización de resoluciones superiores, concretamente la correspondiente a una escala 2 : 1, se simula la utilización de radios inferiores a la unidad con el operador. Con ello, es posible llevar a cabo el análisis de vecindades muy próximas, esperando de esta forma facilitar la segmentación e identificación de los vasos sanguíneos de menor dimensión.

El empleo de la escala 2 : 1 da lugar a cuatro píxeles correspondientes a cada píxel de la resolución 1 : 1, proporcionando de esta forma cuatro características nuevas,

tal y como se detalla en la Sec. 4.2.1.3 y se puede observar en el ejemplo de la Fig. 4.7.

En los resultados mostrados en la Tabla 5.4, estructurados con la misma metodología que los indicados en los experimentos previos, se parte de la escala 1 : 1 como escala única, una combinación con la escala 2 : 1 para la combinación de dos escalas, la adición de la escala 1 : 2 para un número total de escalas igual a 3, y así sucesivamente, considerando sucesivamente resoluciones inferiores. En las Fig. 5.5 y A.6 se proporcionan los resultados gráficos del análisis.

De cara a analizar los resultados proporcionados (véase la Tabla 5.4), se debe considerar la comparación entre la utilización de una única escala, la original, y la utilización de dos escalas simultáneas, 1 : 1 y 2 : 1, donde se observa una mejora relevante de todas las métricas a excepción de la especificidad. Así mismo, se puede establecer una relación con la Tabla 5.2, puesto que, por ejemplo, las 3 escalas empleadas de manera simultánea en dicha tabla (1 : 1, 1 : 2 y 1 : 4) son igualmente utilizadas en la tabla correspondiente a este experimento, junto con la escala 2 : 1, cuando se indica la utilización de un total de 4 escalas. De esta forma, también es posible observar un incremento del poder predictivo, que, sin embargo, se ve reducido a medida que se considera un número mayor de escalas de forma combinada.

Núm. total de escalas	1	2	3	4	5	6
<b>F1-score</b>	0,393	0,477	0,597	0,659	0,677	0,670
<b>Precisión</b>	0,878	0,889	0,907	0,919	0,922	0,919
<b>Sensibilidad</b>	0,306	0,393	0,537	0,607	0,635	0,639
<b>Especificidad</b>	0,963	0,962	0,961	0,965	0,964	0,960

**Tabla 5.4.:** Métricas de evaluación de los resultados de la aplicación de la característica adicional correspondiente a la escala 2 : 1.

### 5.3.5. Selección final de características basadas en el operador *LBP*

Las características generadas mediante los experimentos previos proporcionan, según los resultados obtenidos, mejoras relevantes en los resultados de segmentación sobre el conjunto de imágenes de validación del dataset *DRIVE*. Sin embargo, no se ha analizado detalladamente el impacto sobre los resultados de la utilización conjunta de todas dichas características.

Por lo tanto, se lleva a cabo un análisis de la relevancia de cada uno de los diferentes tipos de características mediante un proceso de validación cruzada en el que se incluyen las 14 imágenes del dataset *DRIVE* correspondientes a la submuestra de entrenamiento, dividiendo las mismas en cuatro conjuntos de tres imágenes y un conjunto adicional de una imagen, con la finalidad de verificar el poder predictivo



**Figura 5.5.:** Resultado gráfico de la aplicación de la característica adicional correspondiente a la escala 2 : 1.

de dichas características, entrenando sobre cuatro de los cinco conjuntos de imágenes y prediciendo sobre el restante un total de cinco veces, de forma que cada uno de los diferentes conjuntos de imágenes sea empleado una vez para la evaluación de los resultados.

La selección final de características se realiza analizando conjuntos de estas provenientes de una metodología común, siendo dichas agrupaciones:

- Escalas: se busca obtener el número de escalas óptimo a emplear, analizándose la relevancia de la utilización de las características a escalas 2 : 1, 1 : 4, 1 : 8 y 1 : 16, las «extremas» de la pirámide de imagen. Dado que se realiza un agrupamiento de características del mismo tipo, el análisis, por ejemplo, de la relevancia de las características a escala 1 : 8 comprende todas las características generadas mediante el reescalado de la imagen original a dicha escala para todas las variantes de los operadores *LBP*.
- Variantes del operador *LBP*: se consideran las cinco variantes, la original, la *RI*, la *RIU2*, la *NRI* y la *VAR*. Dado que el análisis se lleva a cabo de manera agrupada se verifica la relevancia para cada una de sus pirámides de escalas de manera conjunta.
- Intensidad de los píxeles de la imagen original. Se trata de una característica única, descrita en la Sec. 5.3.3.

La selección de características mediante validación cruzada se realiza por medio del descarte de cada una de las agrupaciones de variables mencionadas, manteniendo las

restantes presentes, y el entrenamiento de cinco clasificadores ajustados mediante la técnica de validación cruzada y tomando como métrica ( $F1$ -score) la media de los cinco resultados de evaluación. De esta forma, se compara la métrica del clasificador que emplea la totalidad de características con la de los clasificadores que emplean dichas características menos las correspondientes a la agrupación cuyo poder predictivo se está analizando. Se ha definido un valor límite de 0,005 de disminución del  $F1$ -score para determinar el conjunto de características como poco discriminante y, por lo tanto, descartable en el proceso de selección de características. Se trata de un proceso iterativo en el que se descarta en primer lugar la agrupación cuyo impacto sobre la métrica sea menos perjudicial y, una vez descartado dicho grupo de características, se comprueba el impacto del descarte de las restantes agrupaciones.

En la Tabla 5.5 se encuentran los resultados del proceso. Se parte de un valor de referencia de la métrica  $F1$  de 0,665 obtenido mediante validación cruzada a partir de la totalidad de variables. Como se puede observar, el primer conjunto de características descartado es el correspondiente a la escala 1 : 16, en consonancia con los resultados obtenidos en apartados previos en los que las escalas de menor resolución no producían una mejora en los resultados, presumiblemente, por el proceso de interpolación tan relevante en un reescalado en el que el número de píxeles se ve considerablemente reducido. El siguiente conjunto de características descartado es el correspondiente a la variante del operador  $RIU2$ , la cual simplifica la información proporcionada por el operador  $RI$  mediante la generación de patrones uniformes, disminuyendo la cantidad de información. En la tercera iteración se descartan las características de la variante original del operador, la cual proporciona una mayor cantidad de categorías que la invariante a la rotación, pero dado que para la segmentación de los vasos sanguíneos la orientación de los mismos no debería ser relevante, se puede concluir que la inclusión de la información acerca de la orientación de los patrones en la vecindad de los píxeles no proporciona el suficiente poder predictivo para ser considerada. Finalmente se descarta la variante uniforme dependiente de la rotación,  $NRI$ , donde sería aplicable el criterio concluido en el descarte de la variante original. La métrica  $F1$  finalmente obtenida mediante validación cruzada tras la selección de características es 0,669, superior al valor de partida con todas las características.

Iter.	2:1	1:4	1:8	1:16	$LBP_{def}$ .	RI	RIU2	NRI	VAR	Int. gris
1	0,656	0,653	0,665	<b>0,671</b>	0,666	0,664	0,664	0,664	0,496	0,636
2	0,660	0,660	0,668	-	0,671	0,669	<b>0,672</b>	0,672	0,498	0,639
3	0,660	0,661	0,667	-	<b>0,671</b>	0,670	-	0,671	0,499	0,639
4	0,662	0,658	0,666	-	-	0,666	-	<b>0,669</b>	0,485	0,638
5	0,661	0,658	0,663	-	-	0,576	-	-	0,457	0,635

**Tabla 5.5.:** Resultados de la evaluación mediante la métrica  $F1$  de las iteraciones de selección de características por validación cruzada.

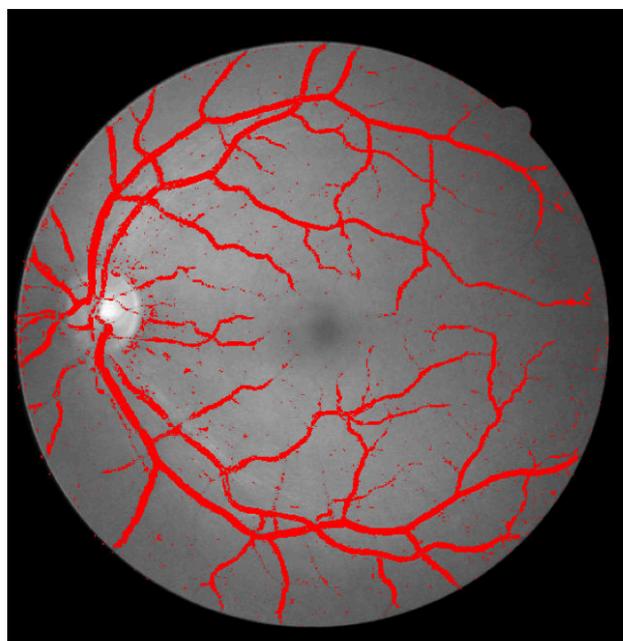
Por lo tanto, dados los resultados obtenidos, la selección final de características

incluye un total de 17 variables, donde 8 variables provienen de la pirámide del operador *LBP* invariante a la rotación (4 de ellas correspondientes a la escala 2 : 1 y las otras 4 correspondientes a las escalas 1 : 1, 1 : 2, 1 : 4 y 1 : 8), 8 de la versión invariante a la rotación, pero dependiente de la intensidad, y, una adicional consistente en la intensidad de los píxeles de la imagen tras el preprocesamiento.

F1-score	Precisión	Sensibilidad	Especificidad
0,705	0,927	0,679	0,963

**Tabla 5.6.:** Métricas de evaluación sobre el conjunto de validación del dataset *DRIVE* basada en un filtrado final del conjunto de características *LBP*.

En la Tabla 5.6 y en la Fig. 5.6 es posible encontrar los resultados de evaluación al predecir sobre el conjunto de imágenes de validación del dataset *DRIVE* tras entrenar el clasificador empleando las imágenes de entrenamiento con las características seleccionadas mediante validación cruzada. Al igual que en el proceso de validación cruzada en el que la utilización conjunta de las características seleccionadas proporciona mejor resultado que el de partida (considerando la métrica *F1* como referencia), en los resultados sobre el conjunto de imágenes de validación se tiene una mejora en el comportamiento predictivo respecto a cualquiera de los resultados obtenidos previamente en la experimentación, confirmando que las características escogidas proporcionan información complementaria entre sí y que en su utilización conjunta proporcionan un elevado poder predictivo.



**Figura 5.6.:** Resultado gráfico de la predicción con las características seleccionadas mediante validación cruzada.

### 5.3.6. Utilización del canal verde

La utilización del canal verde de las imágenes de fondo de ojo en lugar de la intensidad de gris generada por medio de los tres canales *RGB* ha sido ampliamente contrastada en estudios previos, entre los cuales se encuentran [11, 12, 16, 17, 18], proporcionando resultados contrastados mediante diferentes bases de datos. Esto se debe a que proporciona el mayor contraste entre los vasos sanguíneos y el color del fondo, mientras que los canales rojo y azul muestran menos contraste y mayor ruido. Por lo tanto, y dado que los resultados del trabajo se validarán en posteriores apartados con múltiples bases de datos, se procede a sustituir la intensidad de nivel de gris por el empleo del canal verde.

Partiendo únicamente del canal verde de las imágenes y empleando las mismas operaciones de preprocesado que en las operaciones previas, con la selección de características de la Sec. 5.3.5 se obtiene mediante la técnica de validación cruzada un resultado de  $F1 = 0,675$ , considerando igualmente las 14 imágenes definidas como entrenamiento de la base de datos *DRIVE*. Al utilizar la selección de imágenes de validación los resultados son los mostrados en la Tabla 5.7 y en la Fig. 5.7.

F1-score	Precisión	Sensibilidad	Especificidad
0,704	0,930	0,651	0,971

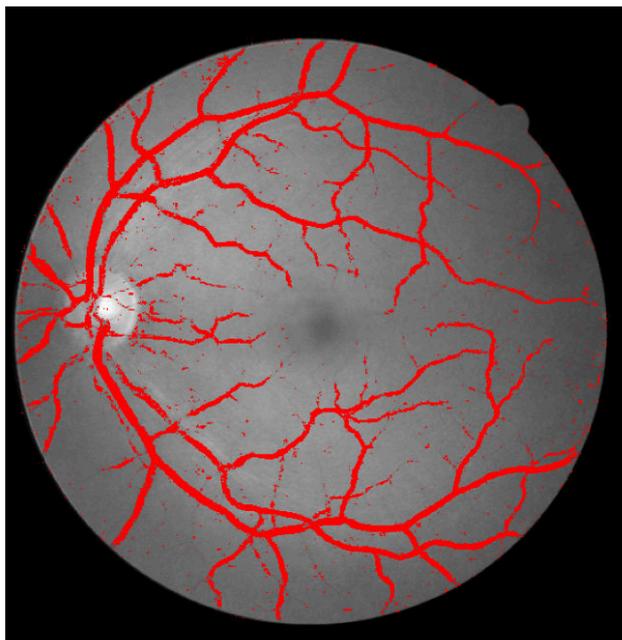
**Tabla 5.7.:** Métricas de evaluación sobre el conjunto de validación del dataset *DRI-VE* partiendo del canal verde.

Dado que los resultados obtenidos mediante validación cruzada son superiores a los obtenidos en las fases previas y en las de validación son equivalentes, junto con el hecho de que su utilización se encuentra recomendada por la literatura y verificada en múltiples bases de imágenes de retina en lugar de únicamente *DRIVE*, como es el caso de este trabajo, se decide emplear dicho canal para todas las etapas de la metodología y de cara a los próximos experimentos.

### 5.3.7. Sintonización de la etapa de preprocesamiento

En los experimentos previos ha sido empleado un preprocesamiento en el cual la sintonización de los parámetros asociados a esta etapa se ha realizado de manera manual en base a los resultados visuales proporcionados en la fase de pruebas. El objetivo de este experimento es realizar una sintonización automática de dichos parámetros mediante computación evolutiva siguiendo los pasos indicados en la Sec. 4.2.3.1.

Los parámetros escogidos de manera automatizada mediante el proceso evolutivo son 41, 6 y 46, para el primer filtro bilateral, correspondientes al diámetro, color y espacio respectivamente, 60 y 32 para el operador *CLAHE*, correspondientes al valor límite y tamaño de la malla, y 32, 42 y 47 para el segundo filtro bilateral.



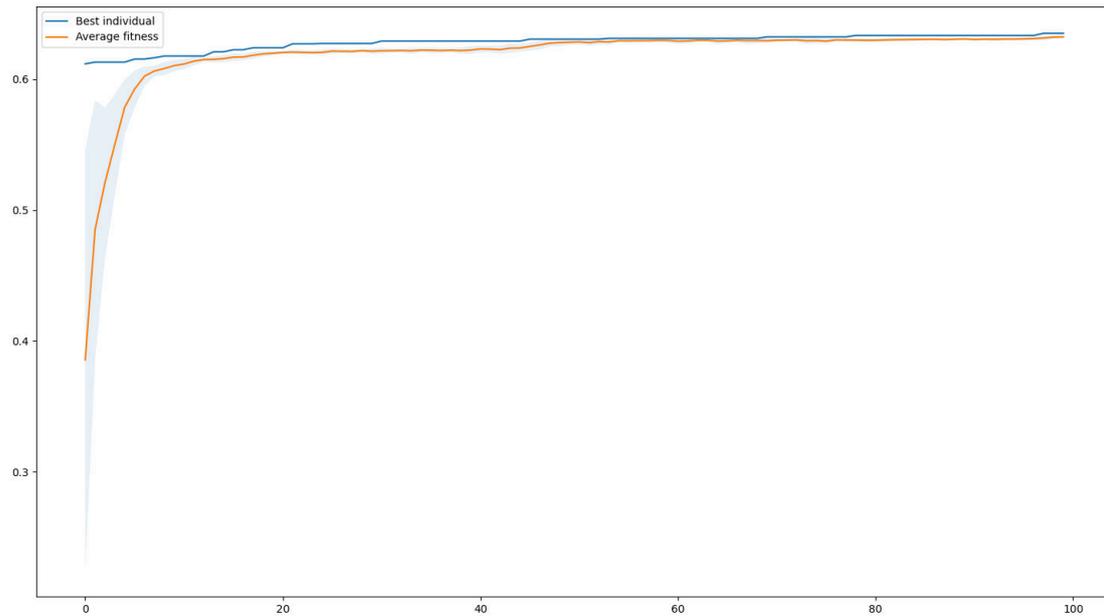
**Figura 5.7.:** Resultado gráfico de la predicción sobre la primera imagen de validación empleando el canal verde.

En la Fig. 4.12 se encuentra un ejemplo de imagen de la base de datos tras la aplicación del preprocesamiento definido. Así mismo, en la Fig. 5.8 se encuentra representada la curva de progreso del resultado medio y del mejor individuo con respecto a la función de evaluación definida para el proceso de optimización a lo largo de las diferentes iteraciones de la ejecución, donde también se representa la varianza de los individuos.

F1-score	Precisión	Sensibilidad	Especificidad
0,717	0,934	0,655	0,975

**Tabla 5.8.:** Métricas de evaluación sobre el conjunto de validación del dataset *DRI-VE* sintonizando automáticamente los parámetros de la etapa de preprocesamiento.

Los resultados obtenidos mediante la sustitución del preprocesamiento manual por el descrito y la selección de características de la Sec. 5.3.5 son de  $F1 = 0,665$  mediante validación cruzada y los indicados en la Tabla 5.8 y la Fig. 5.9. Se observa, en base a los resultados, una disminución de las prestaciones sobre las imágenes de entrenamiento y una mejora sobre las de validación, que se puede deber a la disparidad de las imágenes entre sí, ya que en los experimentos previos ha resultado más sencilla la segmentación de las imágenes de validación puesto que las métricas obtenidas son superiores de forma constante respecto a la submuestra de entrenamiento mediante validación cruzada.



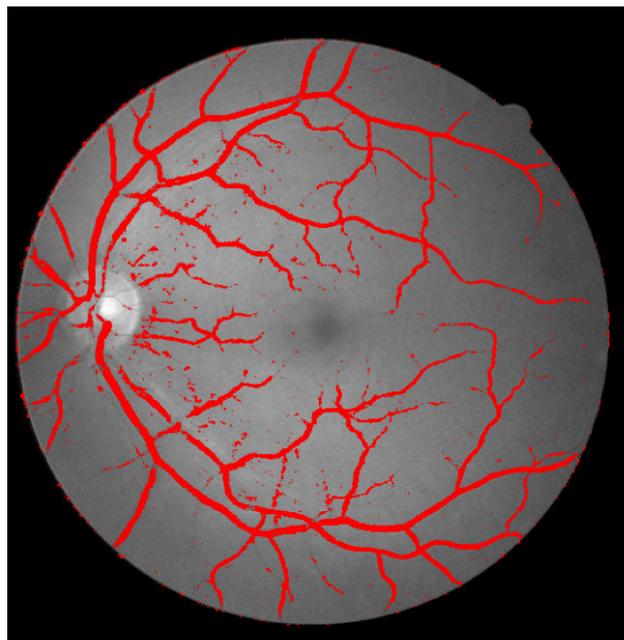
**Figura 5.8.:** Curva de progreso de la evaluación de la población durante el ajuste del preprocesamiento en la base de datos *DRIVE* en la ejecución del *algoritmo evolutivo*.

Por lo tanto, aunque no sea posible definir si se está realizando una mejora de las prestaciones mediante la metodología propuesta, sí que es posible concluir que la utilización del proceso elimina la necesidad de una sintonización manual y, por lo tanto, constituye una mejora de la metodología que facilita así mismo su adaptación a nuevas bases de datos.

### 5.3.8. Generación de características mediante sintonización de núcleos de convolución

Como es indicado en la sección 4.2.3.2, se propone la utilización de seis núcleos de convolución que optimizados de manera conjunta den lugar a seis características que aporten información complementaria, donde cada uno de ellos se corresponde con la aplicación de una convolución sobre el canal verde de la imagen sin aplicación de preprocesamiento. Se espera obtener un mayor poder predictivo al realizar la sintonización de los seis núcleos de manera conjunta mediante un proceso evolutivo similar al utilizado en la sección previa.

En la Fig. 5.10 se encuentra el gráfico de la curva de progreso del proceso evolutivo de optimización para el dataset *DRIVE*. Así mismo, en la Fig. 5.12 se encuentran las imágenes resultantes de la aplicación de los núcleos de convolución obtenidos sobre una de las imágenes de la base de datos. En la primera fila de imágenes se observan los núcleos de dimensión  $3 \times 3$ , en la segunda fila los de dimensión  $5 \times 5$  y en la



**Figura 5.9.:** Resultado gráfico de la predicción empleando el preprocesamiento con parámetros sintonizados mediante un *algoritmo evolutivo*.

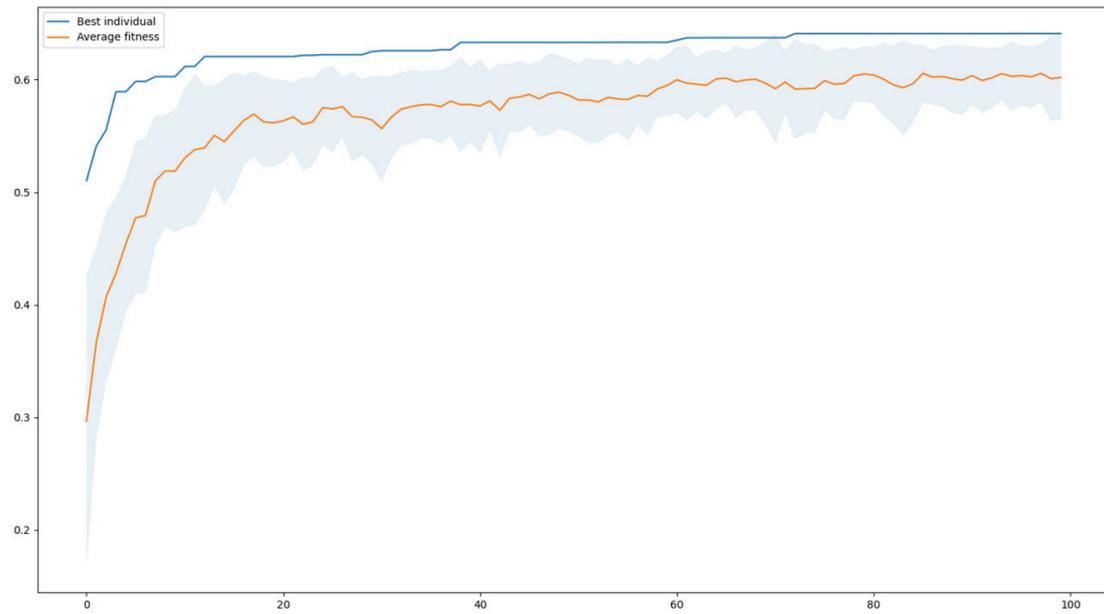
tercera fila las dos imágenes correspondientes a la convolución de los núcleos  $7 \times 7$ .

F1-score	Precisión	Sensibilidad	Especificidad
0,742	0,938	0,697	0,975

**Tabla 5.9.:** Métricas de evaluación sobre el conjunto de validación del dataset *DRI-VE* tras la inclusión de las convoluciones.

En la inclusión de las características generadas mediante el proceso descrito se obtiene un resultado de validación cruzada de  $F1 = 0,699$ . Los resultados obtenidos en el conjunto de imágenes de validación son los indicados en la Tabla 5.9 y en la Fig. 5.9. Se puede observar un incremento relevante del poder predictivo del clasificador ajustado con respecto a los resultados obtenidos previamente en la Tabla 5.8.

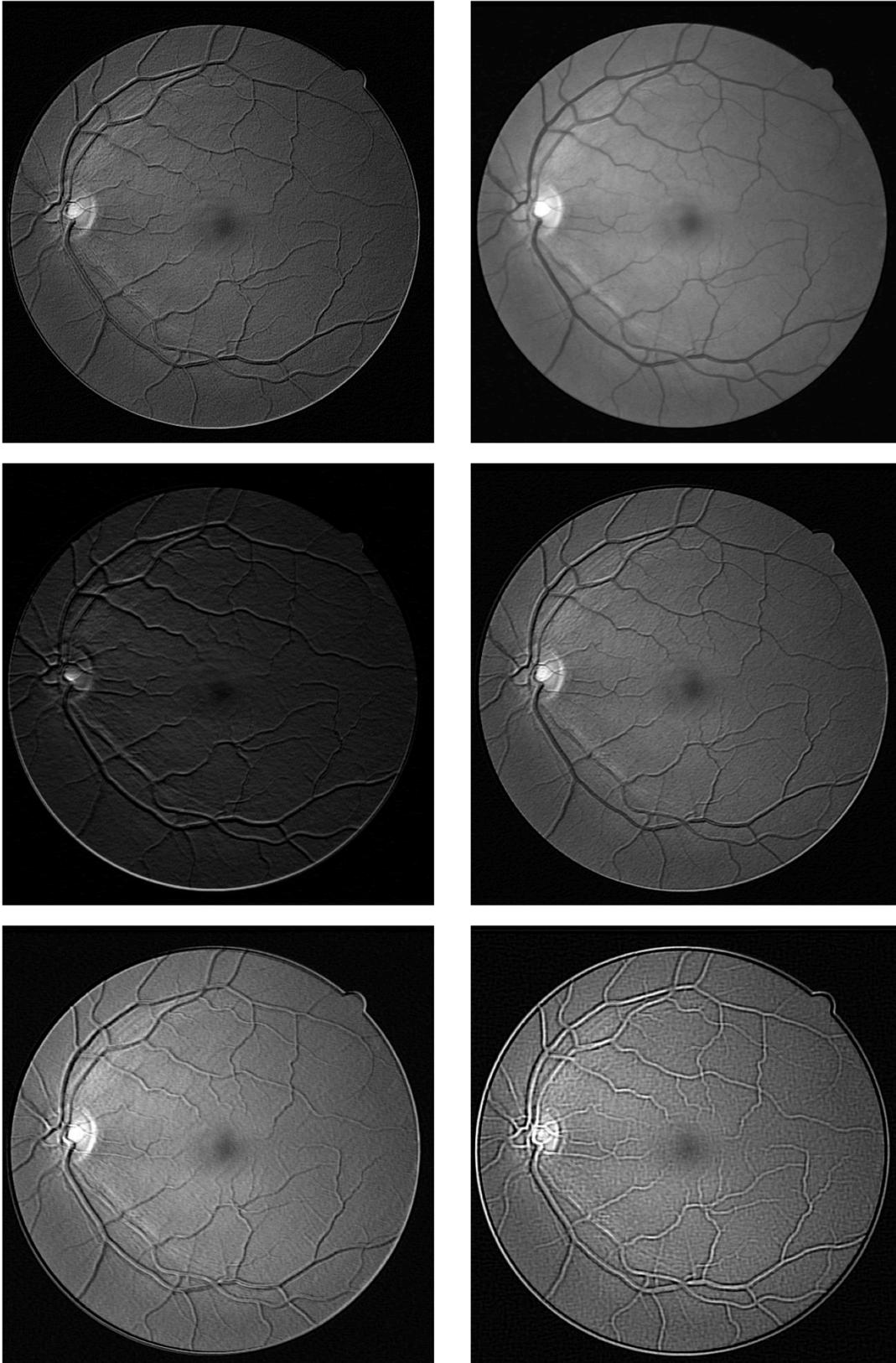
Dados los buenos resultados, el proceso se incluye en la metodología y se lleva a cabo para cada una de las bases de imágenes sobre las que se realizan los procesos de entrenamiento y de evaluación, facilitando así la optimización a las propiedades de las imágenes de cada una de ellas.



**Figura 5.10.:** Curva de progreso de la evaluación de la población durante el ajuste de las convoluciones en la base de datos *DRIVE* en la ejecución del *algoritmo evolutivo*.



**Figura 5.11.:** Resultado gráfico de la predicción sobre la primera imagen de validación de *DRIVE* empleando adicionalmente convoluciones.



**Figura 5.12.:** Representación de las convoluciones sobre una de las imágenes de validación del dataset *DRIVE*.

### 5.3.9. Sintonización de parámetros mediante *Grid Search*

El último experimento en lo relativo a procesos de optimización consiste en la búsqueda de hiperparámetros óptimos del clasificador descrita en la Sec. 4.2.3.3. El proceso iterativo en el que se comprueban todas las posibles combinaciones de hiperparámetros se realiza mediante la siguiente «malla»:

- Número máximo de hojas: 40, 50 y 60.
- Máxima profundidad: 20, 30 y 40.
- Muestreo de las columnas por árbol: 70 % y 90 %.
- Muestreo del número de registros: 50 % y 70 %.
- Ratio de aprendizaje: 0, 2, 0, 5 y 0, 7.

Los hiperparámetros finalmente seleccionados por medio del proceso de optimización descrito son los mostrados en la Tabla 5.10. Los resultados obtenidos sobre el conjunto de datos de validación son los indicados en la Tabla 5.11 y en la Fig. 5.13.

Máx. hojas	Máx. profundidad	Muestro col.	Muestreo reg.	Ratio aprendizaje
40	20	70 %	50 %	0, 2

**Tabla 5.10.:** Selección de hiperparámetros seleccionada para la base de datos *DRI-VE* mediante la técnica *Grid Search*.

Comparando los resultados de la Tabla 5.11 con los obtenidos en el experimento previo (véase la Tabla 5.9) se observa una mejora. El proceso de sintonización de los hiperparámetros será repetido para cada uno de los conjuntos de imágenes sobre los que se aplique la metodología de este trabajo de investigación.

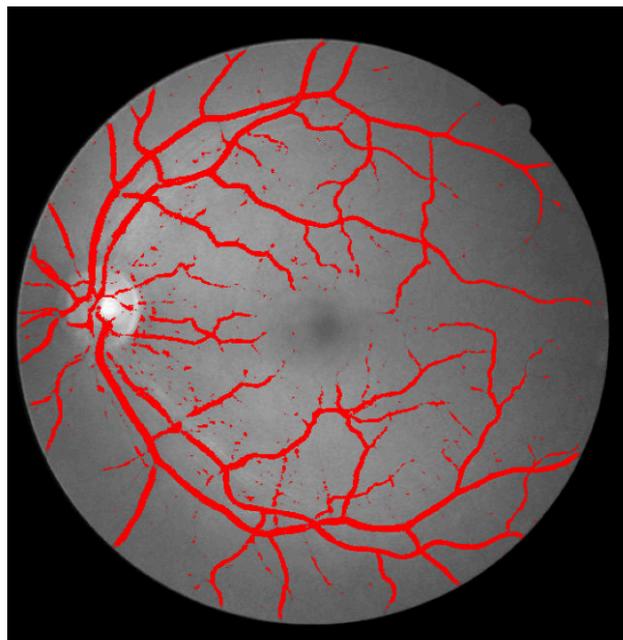
<b>F1-score</b>	<b>Precisión</b>	<b>Sensibilidad</b>	<b>Especificidad</b>
0,749	0,940	0,696	0,976

**Tabla 5.11.:** Métricas de evaluación sobre el conjunto de validación del dataset *DRIVE* en la utilización de parámetros optimizados.

### 5.3.10. Etapa de postprocesamiento

En esta sección se indican los resultados de la realización de la etapa de postprocesamiento en la que se lleva a cabo el entrenamiento de un segundo clasificador que parte de la misma información que el primero junto con resultados predictivos de dicha primera iteración. Con ello, se busca introducir la relación de dependencia existente entre píxeles próximos. El proceso se encuentra descrito en la Sec. 4.2.3.4.

Los resultados obtenidos en el conjunto de imágenes de validación son los de la Tabla 5.12 y Fig. 5.14. Se comprueba un incremento del poder predictivo, siendo



**Figura 5.13.:** Resultado gráfico de la predicción sobre la primera imagen de validación de *DRIVE* en la utilización de parámetros optimizados.

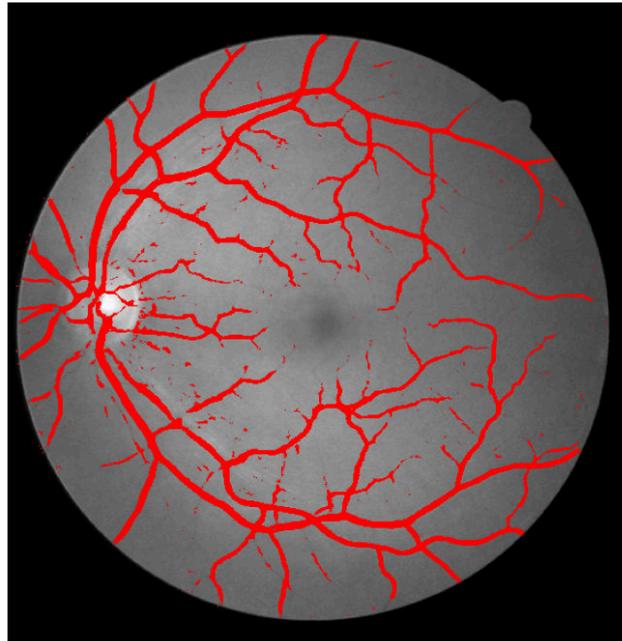
F1-score	Precisión	Sensibilidad	Especificidad
0,765	0,943	0,728	0,974

**Tabla 5.12.:** Métricas de evaluación sobre el conjunto de validación del dataset *DRIVE* tras la etapa de postprocesamiento.

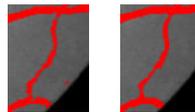
posible así mismo observar una mayor continuidad en las predicciones obtenidas y una reducción de ruido en los resultados gráficos mostrados en la figura con respecto a los de resultados previos sin dicha etapa de postprocesamiento (véase la Fig. 5.13).

Así mismo, es posible visualizar algunas de las mejoras alcanzadas en esta etapa por medio de la comparación de los resultados gráficos obtenidos en la fase previa:

- Obtención de una mayor continuidad en vasos sanguíneos de difícil identificación (Fig. 5.15).
- Reducción del ruido para falsos positivos (véase la Fig. 5.16).
- Corrección de algunas intersecciones entre vasos sanguíneos. La comparación puede ser observada en la Fig. 5.17, donde se visualiza una de las intersecciones entre vasos optimizada gracias al tratamiento, concretamente la correspondiente a los vasos de mayor dimensión. Sin embargo, a pesar de la mejora algunas de las intersecciones restantes no son correctamente rellenadas por píxeles clasificados como vasos sanguíneos.
- Incremento de la continuidad en el nervio óptico (Fig. 5.18).



**Figura 5.14.:** Resultado gráfico de la predicción sobre la primera imagen de validación *DRIVE* tras la aplicación del postprocesamiento.

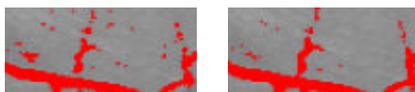


**Figura 5.15.:** Comparación de continuidad en un vaso sanguíneo. La imagen de la derecha incluye el postprocesamiento.

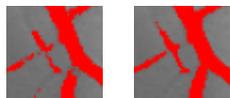
## 5.4. Resultados de evaluación sobre imágenes de test

Desarrollada la experimentación que da lugar a la metodología de segmentación de vasos sanguíneos mediante las técnicas expuestas en la fase de constitución del trabajo de investigación, y con la finalidad de alcanzar los objetivos indicados y de verificar las hipótesis asociadas, se busca proporcionar resultados de test y verificar la posibilidad de aplicación de dicha metodología a nuevos problemas mediante bases de imágenes de retinografías adicionales. De esta forma, aplicando la metodología de la Sec. 4.2 y repitiendo las ejecuciones que precisen una sintonización por medio de la correspondiente división de imágenes en conjuntos de entrenamiento, validación y test, se procede a evaluar sus resultados en las bases de datos *STARE* y *CHASE*, así como en el conjunto de imágenes de test de *DRIVE*.

En la Sec. 4.1 se indican los detalles de las bases de datos en cuanto a resolución, número de retinografías e información adicional como el campo de visión, y en la



**Figura 5.16.:** Comparación relativa a la reducción de ruido. La imagen de la derecha incluye el postprocesamiento.



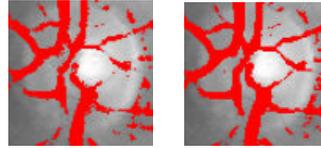
**Figura 5.17.:** Comparación de las intersecciones entre vasos. La imagen de la derecha incluye el postprocesamiento.

Sec. 5.1 se precisan las divisiones en conjuntos de datos de entrenamiento, validación y test, «estandarizadas» mediante las investigaciones previas de la literatura, conservando siempre una selección de imágenes mantenida de forma independiente de todos los procesos previos a la evaluación.

Por lo tanto, para *CHASE* se tiene un proceso similar al expuesto para el dataset *DRIVE* con una única división de las imágenes, y, sin embargo, en *STARE* se emplea un conjunto de entrenamiento de 19 imágenes con una única imagen de test, dividiendo el de entrenamiento a su vez en 13 y 6 imágenes para el propio entrenamiento y para validación, y repitiendo las ejecuciones de la metodología definida 20 veces, es decir, tantas ejecuciones como el número de imágenes que componen el dataset, empleando una imagen diferente como test en cada uno de dichos procesos (metodología *leave one out*).

Con objetivo de proporcionar los resultados de la manera más completa posible, se proporcionan los mismos tanto de forma genérica a la selección de imágenes de evaluación como individualizados a cada una de ellas, por medio tanto de los valores numéricos de las métricas como por una representación de la clasificación de los píxeles sobre cada una de las imágenes. Las métricas se obtienen únicamente sobre los píxeles que forman parte del campo de visión tras la aplicación de las correspondientes máscaras.

En la Tabla 5.13 se muestra un resumen de los resultados recogidos en el Cap. 2 correspondiente al *Estado de la Cuestión* a partir de artículos que emplean las mismas bases de datos con la finalidad de facilitar una comparación con los resultados obtenidos en este trabajo de investigación. Dado que, tal y como ha sido comentado en la introducción de este capítulo, no todos los análisis han sido evaluados en las mismas condiciones en lo que a selección de imágenes se refiere, se han marcado sobre dicha tabla aquellos que han sido obtenidos mediante un conjunto de evaluación igual al de esta investigación. Las métricas restantes han sido incluidas con la finalidad de facilitar un orden de magnitud de manera informativa. Por último, se debe tener en cuenta que no es objeto de este trabajo obtener unos resultados competitivos con las técnicas de *Deep Learning*.



**Figura 5.18.:** Comparación del nervio óptico. La imagen de la derecha incluye el postprocesamiento.

		No superv.		Supervisado					
		-		<i>Deep Learning</i>			Clasif. tradicionales		
		[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]
DRIVE	F1	-	-	<b>0,838</b>	0,829	<b>0,832</b>	-	-	-
	Sen.	0,748	0,784	<b>0,889</b>	0,825	<b>0,833</b>	<b>0,755</b>	<b>0,714</b>	0,786
	Esp.	0,975	0,979	<b>0,977</b>	0,976	<b>0,984</b>	<b>0,980</b>	<b>0,987</b>	0,973
	Prec.	0,960	0,951	<b>0,969</b>	0,957	<b>0,971</b>	<b>0,951</b>	<b>0,961</b>	0,958
CHASE	F1	-	-	<b>0,814</b>	0,828	<b>0,824</b>	-	-	-
	Sen.	-	-	<b>0,837</b>	0,820	<b>0,845</b>	-	-	-
	Esp.	-	-	<b>0,985</b>	0,983	<b>0,986</b>	-	-	-
	Prec.	-	-	<b>0,976</b>	0,983	<b>0,977</b>	-	-	-
STARE	F1	-	-	<b>0,844</b>	-	<b>0,853</b>	-	-	-
	Sen.	0,730	0,753	<b>0,829</b>	-	<b>0,852</b>	0,756	-	0,831
	Esp.	0,983	0,981	<b>0,989</b>	-	<b>0,988</b>	0,984	-	0,963
	Prec.	0,953	0,952	<b>0,977</b>	-	<b>0,978</b>	0,961	-	0,953

**Tabla 5.13.:** Recopilación de las métricas obtenidas en la literatura. Se han resaltado en negrita aquellos valores en los que la evaluación se ha llevado a cabo en las mismas condiciones que en este trabajo.

### 5.4.1. Base de datos *DRIVE*

En la Tabla 5.14 y en las Fig. 5.19, 5.20, 5.21 y 5.22 se proporcionan los resultados individuales de las imágenes de test, mientras que en las Tablas 5.15 y 5.16 se proporcionan las métricas de la consideración del conjunto global de píxeles que forman parte del campo de visión de las retinografías.

	F1-score	Sensibilidad	Especificidad	Precisión
01_test.tif	0,8031	0,7913	0,9729	0,9491
02_test.tif	0,8301	0,7888	0,9803	0,9516
03_test.tif	0,7419	0,6422	0,9848	0,9345
04_test.tif	0,7934	0,7252	0,9842	0,9496
05_test.tif	0,7517	0,6391	0,9904	0,9427
06_test.tif	0,7352	0,6243	0,9878	0,9365
07_test.tif	0,7718	0,6905	0,9849	0,9459
08_test.tif	0,7129	0,6014	0,9877	0,9391
09_test.tif	0,6896	0,5522	0,9934	0,9416
10_test.tif	0,7549	0,6787	0,9838	0,9474
11_test.tif	0,7789	0,7451	0,9750	0,9452
12_test.tif	0,7753	0,7231	0,9800	0,9475
13_test.tif	0,7440	0,6358	0,9879	0,9380
14_test.tif	0,7815	0,7835	0,9704	0,9483
15_test.tif	0,7192	0,7941	0,9520	0,9356
16_test.tif	0,7854	0,7157	0,9839	0,9488
17_test.tif	0,7100	0,5929	0,9891	0,9403
18_test.tif	0,7638	0,7234	0,9778	0,9486
19_test.tif	0,8298	0,8213	0,9783	0,9594
20_test.tif	0,7830	0,7458	0,9810	0,9559

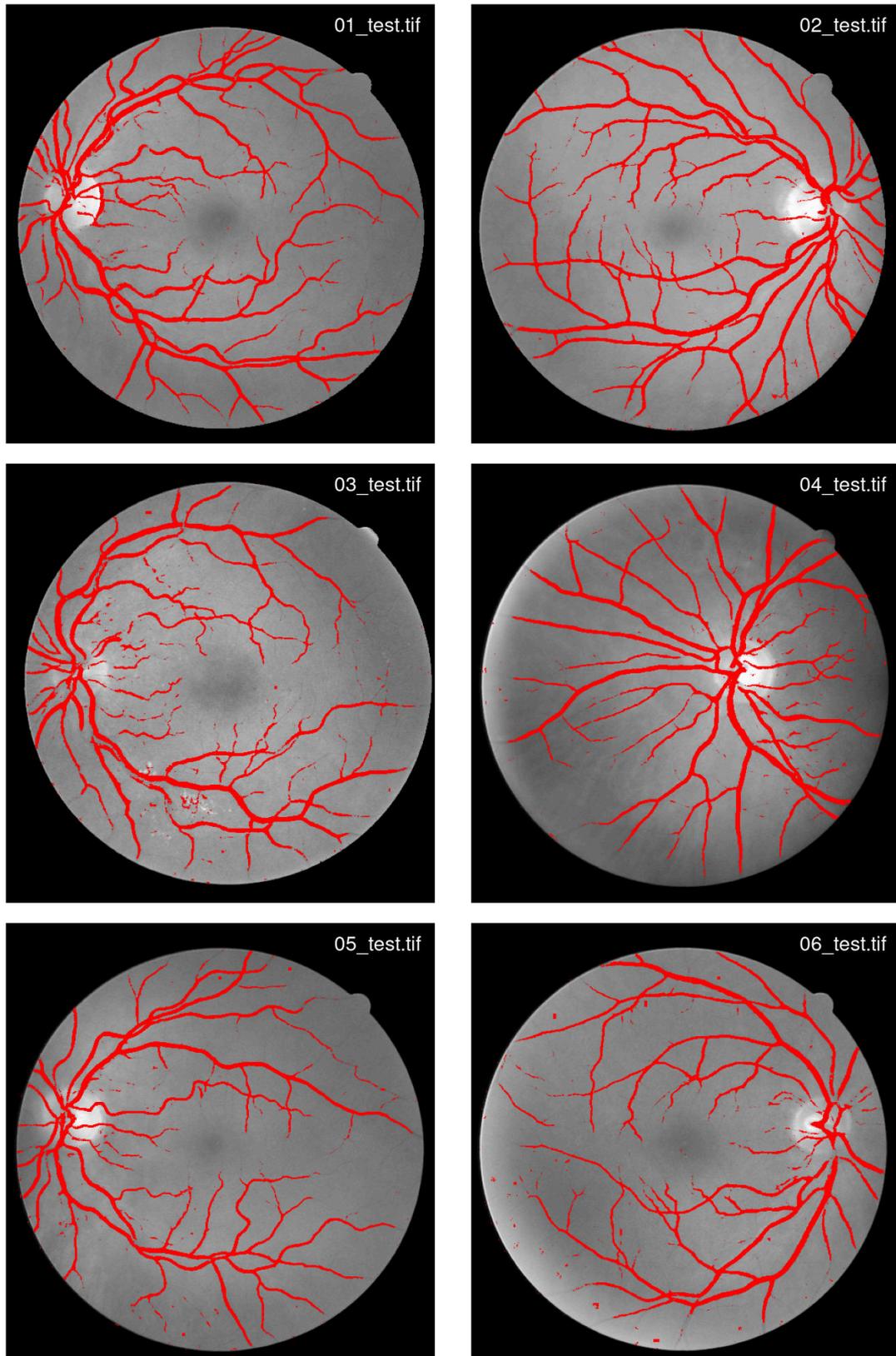
**Tabla 5.14.:** Métricas obtenidas en la evaluación de cada una de las imágenes de test de *DRIVE*.

Real/Pred.	Negativo	Positivo
Negativo	3886102	74392
Positivo	173777	403872

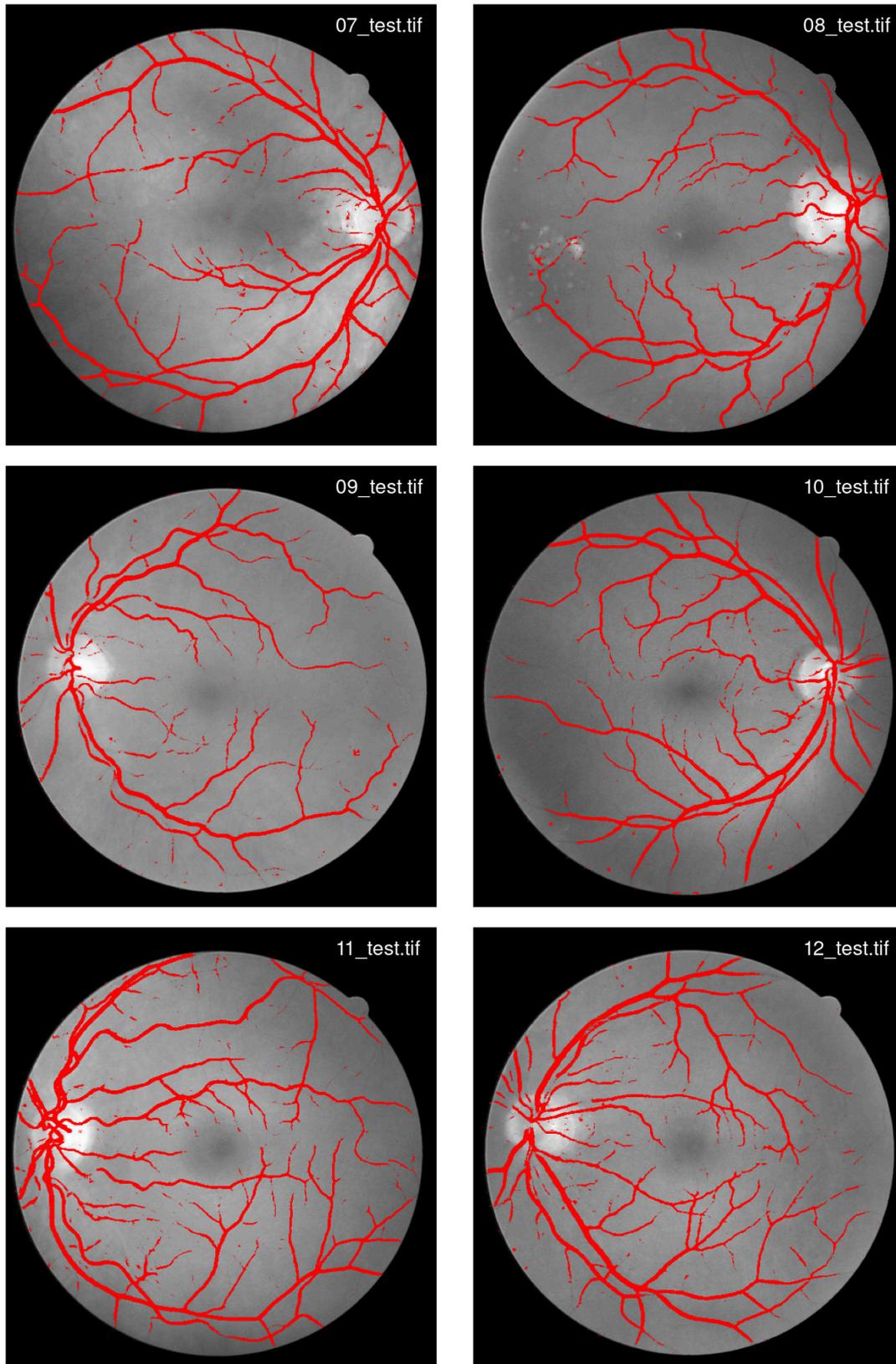
**Tabla 5.15.:** Matriz de confusión resultante en la evaluación del conjunto de datos de test de *DRIVE*.

F1-score	Sensibilidad	Especificidad	Precisión
0,7650	0,6992	0,9812	0,9453

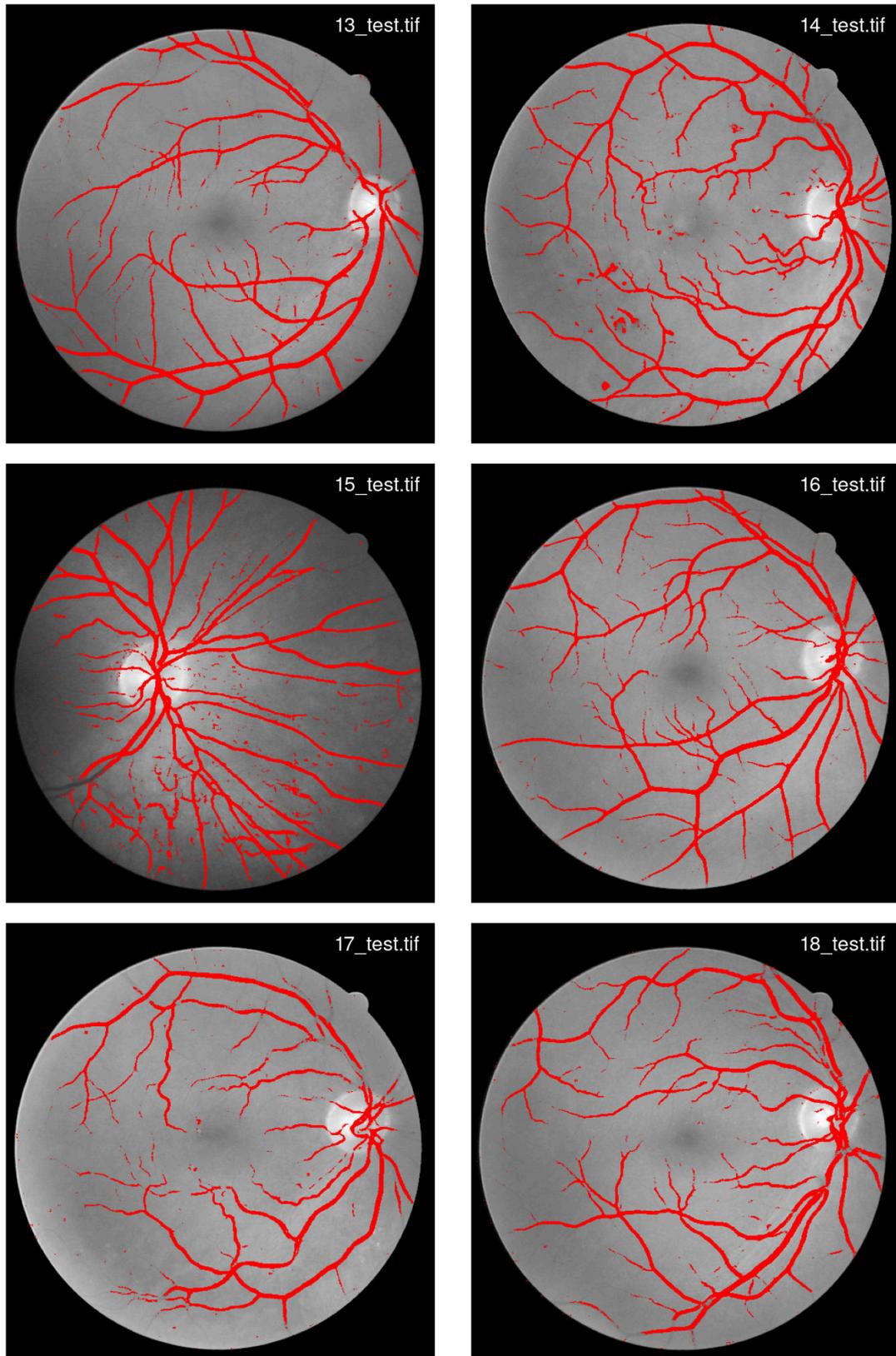
**Tabla 5.16.:** Métricas de la consideración global en la evaluación de todos los píxeles de imágenes de test de *DRIVE*.



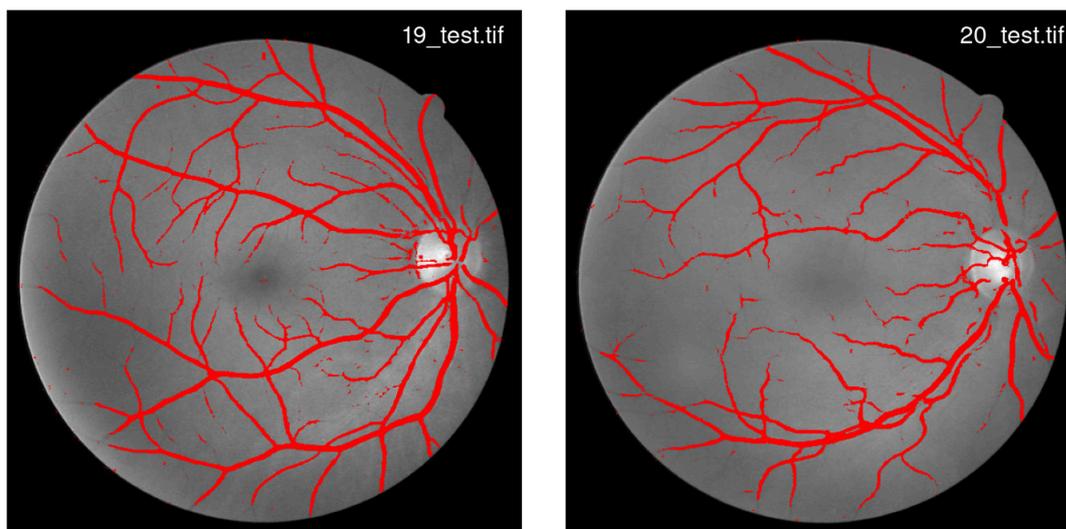
**Figura 5.19.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *DRIVE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



**Figura 5.20.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *DRIVE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



**Figura 5.21.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *DRIVE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



**Figura 5.22.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *DRIVE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.

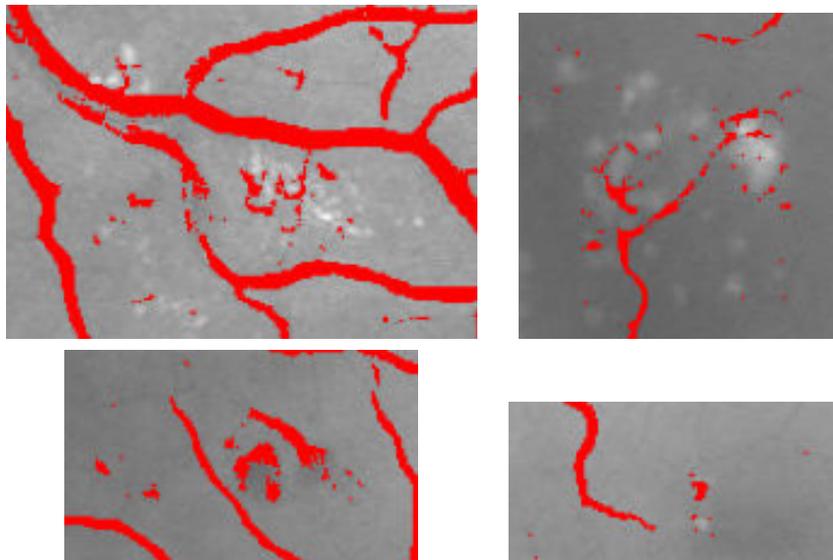
Revisando las métricas proporcionadas por cada una de las imágenes, en lo relativo a la métrica precisión se observa un valor similar entre las mismas, con variaciones aproximadamente de un 2% de los posibles valores que puede adquirir, mientras que las restantes métricas sí que muestran una mayor varianza. En lo relativo a la especificidad, se tiene un margen de un 4%, que, dado que únicamente el 11,78% de los píxeles de las 20 imágenes se corresponden con vasos sanguíneos, es una elevada variación. Dicha varianza es más sencilla de observar en la métrica de sensibilidad (27%). En lo relativo al *F1-score*, que refleja un «equilibrio» entre las métricas de sensibilidad y especificidad, se tiene una variación del 14%, donde se considera, por lo tanto, que dicha métrica es la más representativa, confirmando la relevancia de su utilización en las fases de experimentación.

Una de las principales discusiones que se puede extraer respecto al análisis visual de las predicciones de las imágenes es la relevancia del valor medio de intensidad de la retinografía, es decir, la «oscuridad» de la imagen. En aquellas con un menor valor de sensibilidad, como *09\_test* o *17\_test*, se observan tonalidades más claras, mientras que en las que tienen una sensibilidad más elevada, como en *15\_test*, donde hay intensidades más oscuras. Esto se puede deber a que los vasos sanguíneos tienen por lo general unas tonalidades más oscuras y, por lo tanto, que se produzca un sesgo hacia dichos valores.

Respecto a los resultados obtenidos en investigaciones estudiadas en el *Estado de la Cuestión*, las indicadas en la Tabla 5.13, se obtienen peores métricas que en los trabajos supervisados, alcanzando en el caso de los basados en metodologías que no emplean *Deep Learning* una especificidad similar, pero obteniendo en este trabajo una sensibilidad algo inferior. En las dos investigaciones basadas en clasificadores

tradicionales que emplean la misma metodología de evaluación que la expuesta en este trabajo, [16, 17], es posible estimar el valor  $F1$ -score por medio de las métricas proporcionadas y el número total de píxeles del conjunto de datos de test, proporcionando valores  $F1 = 0,799$  [16] y  $F1 = 0,777$  [17], resultados son directamente comparables con los de este trabajo de investigación, donde se ha obtenido un valor  $F1 = 0,765$ , y, por lo tanto, se puede considerar que se trata de resultados competitivos.

En lo relativo imágenes con patologías, presentes en las imágenes  $03\_test$ ,  $08\_test$ ,  $14\_test$  y  $17\_test$ , se tiene una disminución de las prestaciones que puede ser observada visualmente en algunas regiones (véase la Fig. 5.23). Sin embargo, considerando las métricas obtenidas en la Tabla 5.14 el impacto es moderado. Dichas métricas no se diferencian de los valores globales de la Tabla 5.16 en mayor medida que otras imágenes, llegando a superar dichos valores en la imagen  $14\_test$ , la cual posee uno de los impactos más importantes debido a patologías, etiquetándose algunas de dichas regiones como vasos sanguíneos, como puede ser observado visualmente.



**Figura 5.23.:** Regiones con patologías de las imágenes de *DRIVE*. Se tiene de color rojo aquellos píxeles etiquetados como vasos sanguíneos en el proceso de predicción.

#### 5.4.2. Base de datos *CHASE*

En la Tabla 5.17 y en las Fig. 5.24 y 5.25, se muestran los resultados individuales de las imágenes de test, mientras que en las Tablas 5.18 y 5.19 se proporcionan las métricas de la consideración del conjunto global de píxeles que forman parte del campo de visión de las retinografías.

	F1-score	Sensibilidad	Especificidad	Precisión
Image_11L.jpg	0,7364	0,7207	0,9804	0,9605
Image_11R.jpg	0,7227	0,6698	0,9848	0,9606
Image_12L.jpg	0,6972	0,6803	0,9685	0,9386
Image_12R.jpg	0,7485	0,7418	0,9715	0,9472
Image_13L.jpg	0,6752	0,5936	0,9839	0,9492
Image_13R.jpg	0,6425	0,5684	0,9798	0,9421
Image_14L.jpg	0,7396	0,6677	0,9849	0,9535
Image_14R.jpg	0,6952	0,6934	0,9722	0,9486

**Tabla 5.17.:** Métricas obtenidas en la evaluación de cada una de las imágenes de test de *CHASE*.

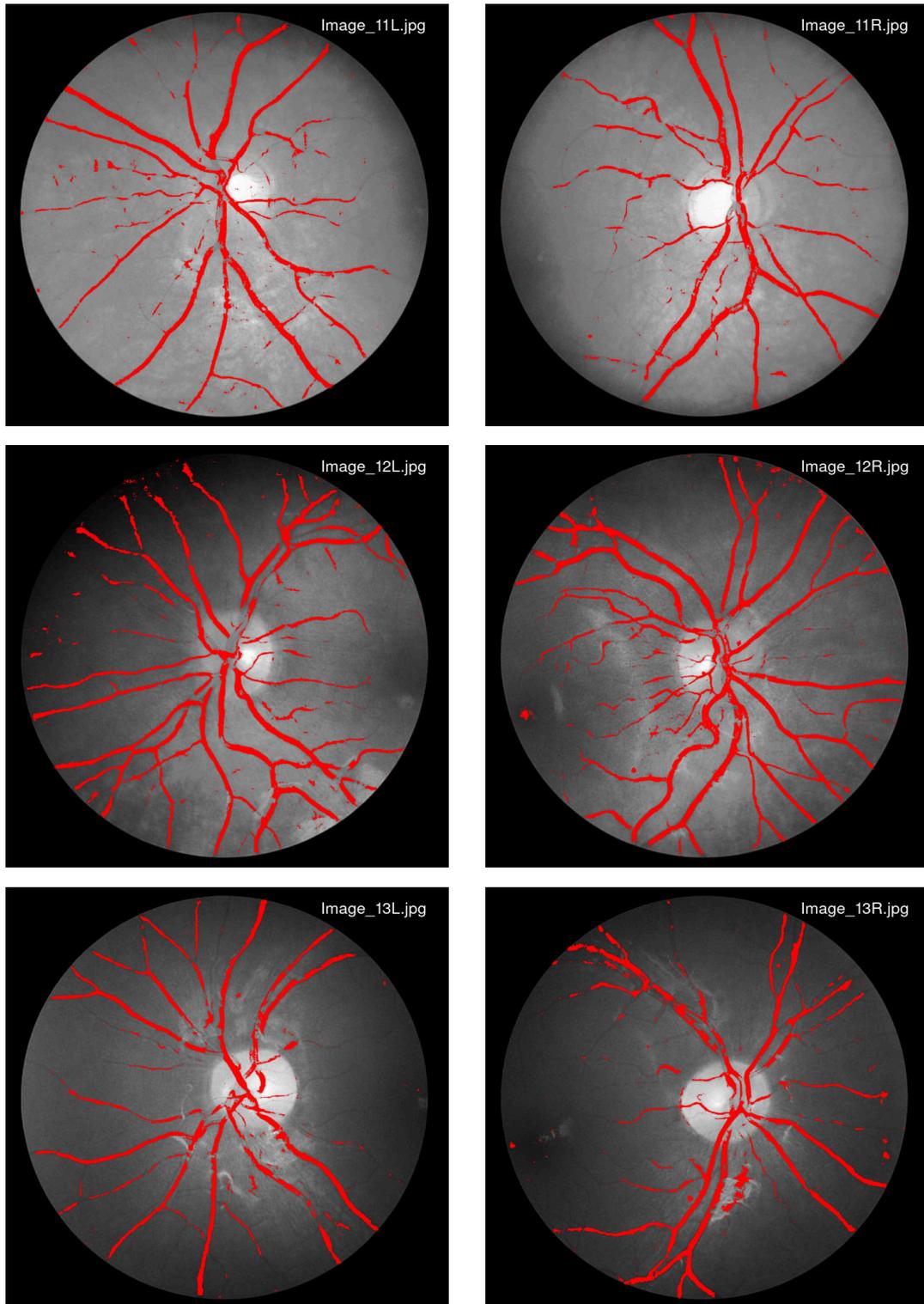
Real/Pred.	Negativo	Positivo
Negativo	4731063	105038
Positivo	160660	322727

**Tabla 5.18.:** Matriz de confusión resultante en la evaluación del conjunto de datos de test de *CHASE*.

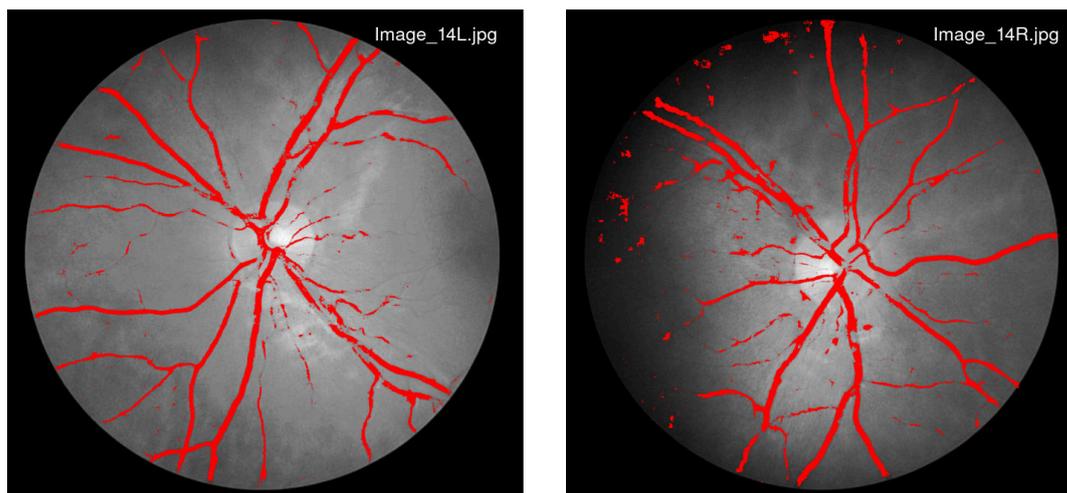
F1-score	Sensibilidad	Especificidad	Precisión
0,7084	0,6676	0,9783	0,9501

**Tabla 5.19.:** Métricas de la consideración global en la evaluación de todos los píxeles de imágenes de test de *CHASE*.

Dados los resultados obtenidos en la base de imágenes de *CHASE*, considerando tanto los resultados de la representación gráfica como las métricas de evaluación de cada una de las imágenes, se puede apreciar una de las consideraciones discutidas en la sección de resultados de test de *DRIVE* (véase la Sec. 5.4.1), la importancia de la intensidad media de las retinografías. Las que muestran tonos más oscuros, junto aquellas con un menor contraste con los límites de la captura (imágenes *Image\_12L*, *Image\_13L*, *Image\_13R* y *Image\_14R*) son las que precisamente poseen una métrica *F1-score* inferior al valor medio. Sin embargo, no se observa el efecto de un incremento de la sensibilidad en aquellas con tonalidades más oscuras como se apreció para el dataset *DRIVE*.



**Figura 5.24.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *CHASE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



**Figura 5.25.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *CHASE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.

Aunque la muestra de imágenes de test es inferior, y, por lo tanto, más complejo obtener información sobre la varianza existente en los resultados en imágenes con propiedades ligeramente diferentes, se obtienen valores elevados y similares entre sí de especificidad y una variación de hasta el 14% en el rango de valores que puede alcanzar la métrica *F1-score* (similar a *DRIVE*), derivada de la alta varianza de la sensibilidad, es decir, del porcentaje de píxeles de vasos sanguíneos correctamente etiquetados en el proceso predictivo.

Respecto a los resultados existentes en la literatura (véase la Tabla 5.13), únicamente es posible establecer una comparación con investigaciones en las que se ha trabajado con *Deep Learning* puesto que son las únicas que emplearon esta base de datos. Por lo tanto, como es de esperar, los resultados obtenidos mediante redes neuronales convolucionales son superiores, contando concretamente con una sensibilidad muy elevada cuando se compara con los resultados de este trabajo de investigación.

Dado que no se aportan datos acerca de las patologías en la documentación de la base de datos no se realiza un análisis sobre este aspecto en esta sección.

### 5.4.3. Base de datos *STARE*

En la Tabla 5.20 y en las Fig. 5.26, 5.27, 5.28 y 5.29 se proporcionan los resultados individuales de las imágenes de test. Así mismo, en la Tabla 5.21 se indican las métricas medias de la Tabla 5.20. Dada la metodología *leave one out* empleada para la evaluación de los resultados, en lugar de una agregación de los resultados de todos los píxeles se proporciona de forma media puesto que se trata de procesos

independientes de predicción, cada uno con sus particularidades, y, por lo tanto, carece de sentido la consideración de todos los píxeles de forma conjunta.

	F1-score	Sensibilidad	Especificidad	Precisión
im0001.ppm	0,6337	0,5590	0,9751	0,9300
im0002.ppm	0,6257	0,6164	0,9647	0,9332
im0003.ppm	0,6920	0,7804	0,9582	0,9438
im0004.ppm	0,3782	0,2569	0,9885	0,9142
im0005.ppm	0,6879	0,6577	0,9646	0,9271
im0044.ppm	0,7301	0,7832	0,9657	0,9498
im0077.ppm	0,8400	0,8851	0,9729	0,9633
im0081.ppm	0,8010	0,8818	0,9639	0,9556
im0082.ppm	0,8435	0,8248	0,9844	0,9674
im0139.ppm	0,7058	0,8586	0,9295	0,9218
im0162.ppm	0,8008	0,8153	0,9763	0,9608
im0163.ppm	0,8487	0,8572	0,9810	0,9680
im0235.ppm	0,8256	0,7810	0,9847	0,9601
im0236.ppm	0,8168	0,7535	0,9871	0,9584
im0239.ppm	0,6476	0,5090	0,9917	0,9353
im0240.ppm	0,6696	0,5359	0,9895	0,9263
im0255.ppm	0,8138	0,7755	0,9820	0,9569
im0291.ppm	0,6123	0,4662	0,9958	0,9594
im0319.ppm	0,5461	0,3912	0,9974	0,9620
im0324.ppm	0,4019	0,2622	0,9958	0,9295

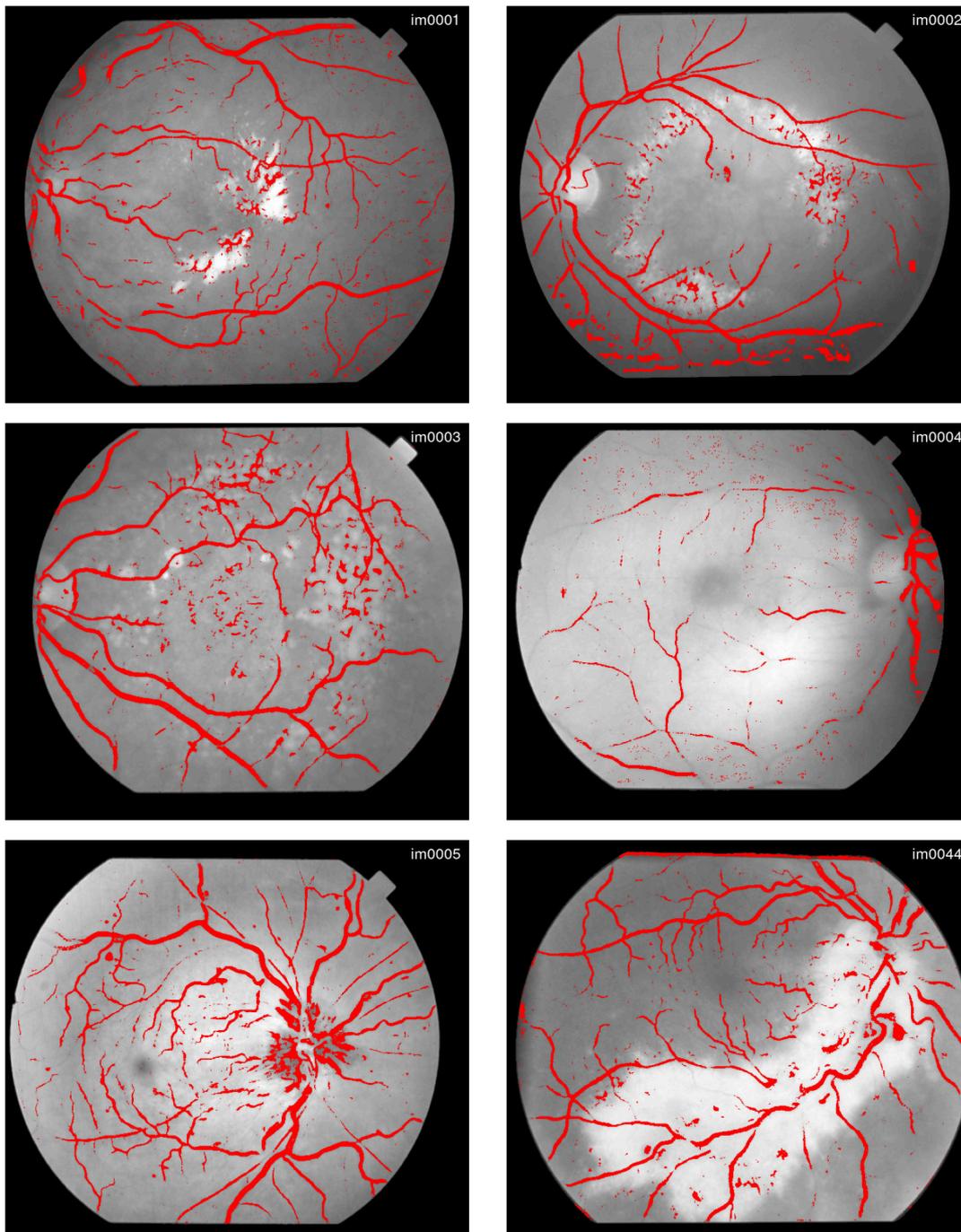
**Tabla 5.20.:** Métricas obtenidas en la evaluación de cada una de las imágenes de test de *STARE*.

F1-score	Sensibilidad	Especificidad	Precisión
0,6961	0,6625	0,9774	0,9461

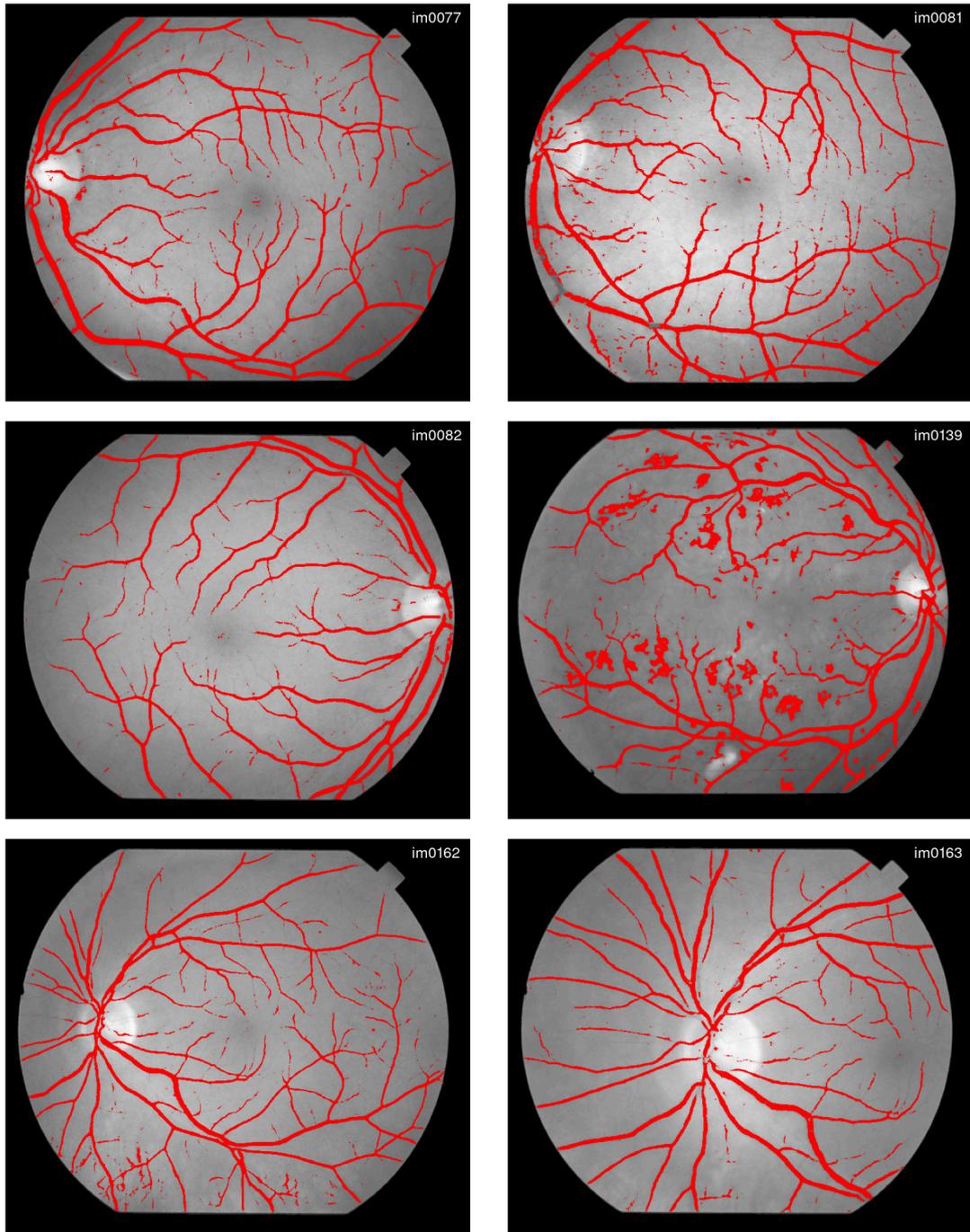
**Tabla 5.21.:** Métricas medias de las imágenes de test de *STARE*. Los valores medios han sido calculados por medio de cada uno de los valores individuales de la Tabla 5.20.

En los resultados alcanzados se producen comportamientos dispares para cada una de las retinografías, con valores de sensibilidad que varían entre 0,25 y 0,88. En base a dicha disparidad, es posible concluir que la técnica *leave one out*, que realiza una validación cruzada «extrema» en la que hay tantos *folds* como imágenes, es una aproximación correcta puesto que una selección aleatoria de imágenes como conjunto de test puede proporcionar resultados muy dispares. Adicionalmente, la

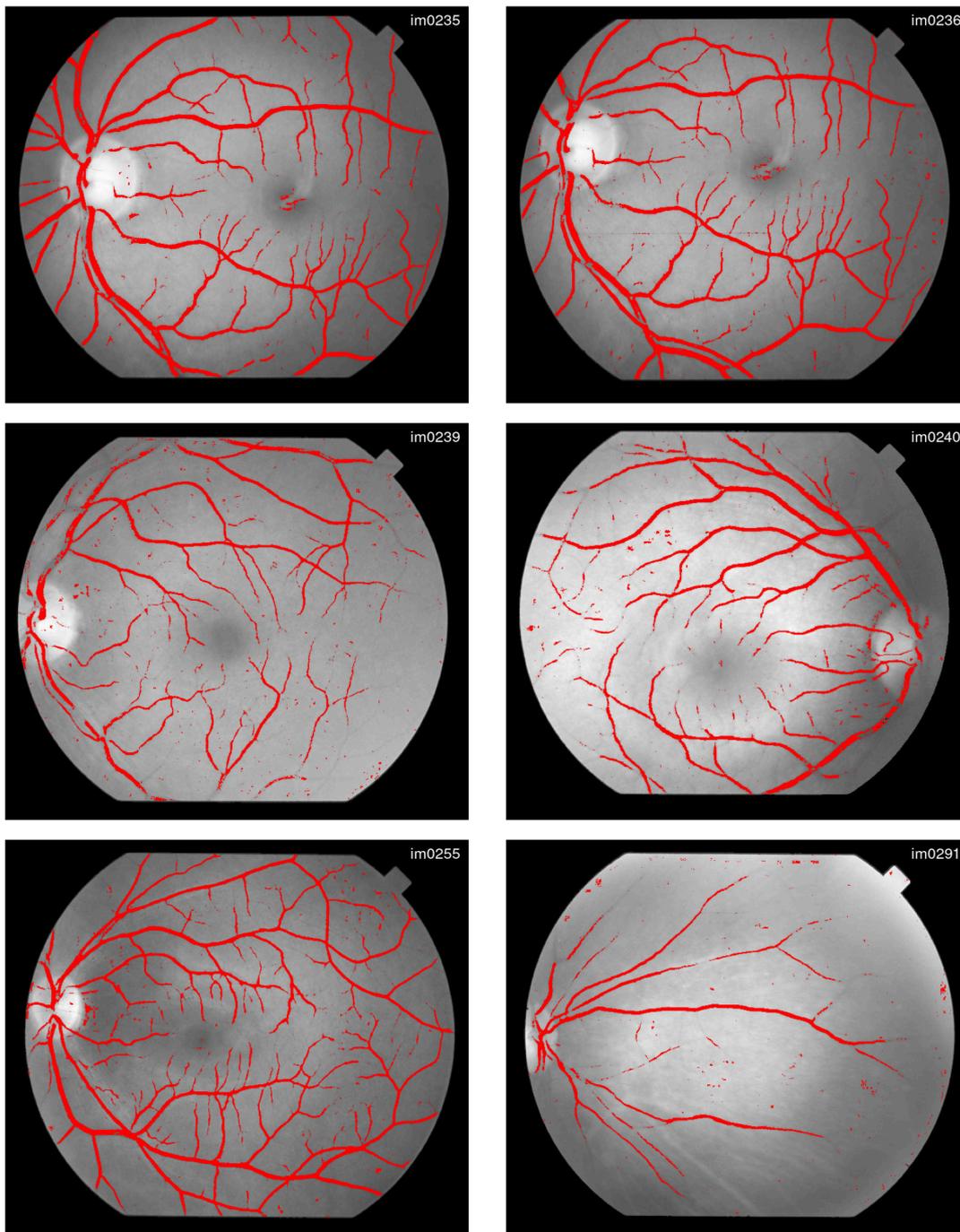
disponibilidad de un mayor número de imágenes para el entrenamiento permite aprovechar en mayor medida el limitado número de imágenes.



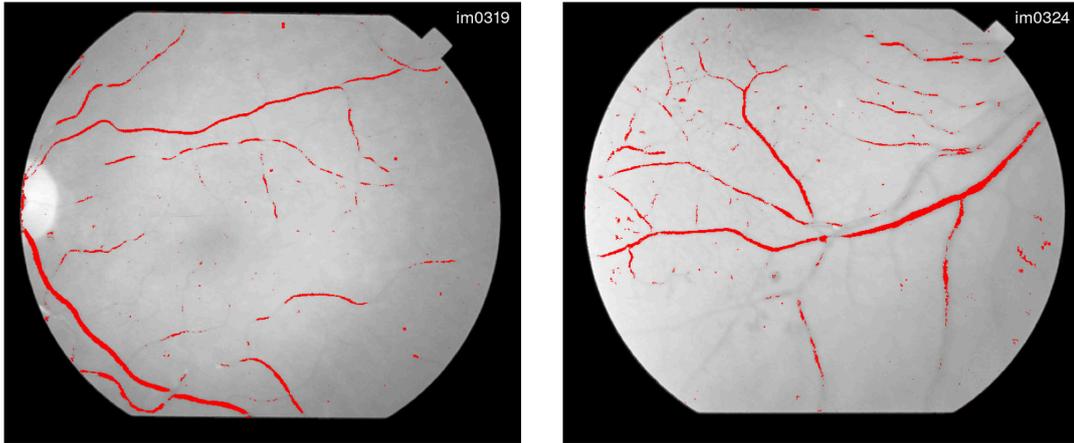
**Figura 5.26.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *STARE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



**Figura 5.27.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *STARE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



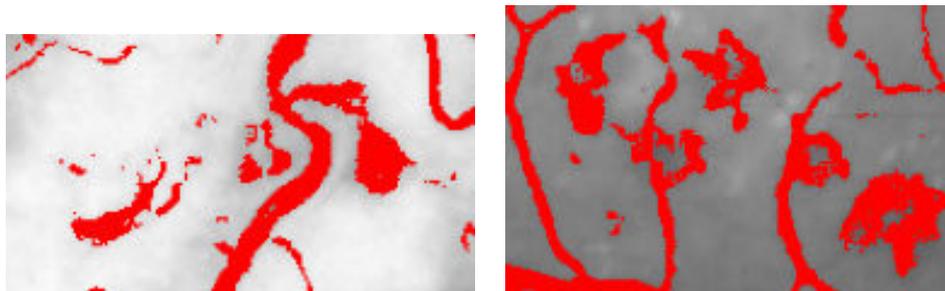
**Figura 5.28.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *STARE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.



**Figura 5.29.:** Representación de las predicciones de los vasos sanguíneos en el conjunto de test de *STARE* sobre las imágenes originales. Los píxeles etiquetados como vasos se encuentran representados en color rojo.

Así mismo, en algunas de las imágenes (*im0004*, *im0291*, *im0319* y *im0324*) se observa cómo resulta complejo mediante un análisis simplemente visual la identificación de los vasos sanguíneos puesto que muestran un menor contraste entre los vasos y sus regiones circundantes. En ellas se muestra un peor resultado de test, con una sensibilidad muy inferior a la obtenida en otras imágenes.

Las patologías se encuentran presentes en las imágenes *im0001*, *im0002*, *im0003*, *im0004*, *im0005*, *im0044*, *im0139*, *im0291*, *im0319* y *im0324*. En ellas se observa un comportamiento dispar en los resultados de evaluación, donde únicamente *im0044* y *im0139* proporcionan unos resultados de clasificación (considerando la métrica *F1-score*) superiores a los del valor medio de la Tabla 5.21. En ambas imágenes se puede identificar mediante un análisis visual sobre los resultados predictivos de las Fig. 5.26 y 5.27 el efecto negativo sobre la capacidad discriminatoria debido a las patologías. Concretamente se tienen algunas regiones con patologías con predicciones incorrectas, como se puede observar en mayor detalle en la Fig. 5.30.



**Figura 5.30.:** Regiones de las imágenes *im0044* y *im0139* (izquierda y derecha respectivamente) sobre las que se observa un empeoramiento del poder predictivo debido a las patologías presentes.

## 6. Conclusiones y trabajo futuro

En este capítulo se encuentran las principales conclusiones extraídas a partir de la investigación y experimentación realizada durante el trabajo, verificando así el cumplimiento de los objetivos propuestos en la fase inicial de definición de este trabajo de investigación (véase la Sec. 1.4). Adicionalmente, por medio de los resultados alcanzados en las diferentes bases de datos, se extraen conclusiones secundarias que resultan de interés, definiendo en base a ellas una serie de líneas para la continuación de la investigación y para la mejora de las principales debilidades encontradas por medio de los resultados. Se trata de líneas que debido al alcance limitado del trabajo no han podido ser abordadas pero que se espera que sean capaces de proporcionar una segunda versión de la metodología que optimice la solución del problema planteado.

### 6.1. Conclusiones

En base a los experimentos realizados con la finalidad de desarrollar y definir la metodología de solución al problema planteado en el trabajo de investigación, se pueden extraer las siguientes conclusiones relativas a las hipótesis y objetivos planteados en este trabajo:

- La hipótesis principal del trabajo de investigación, consistente en la equivalencia entre el operador *LBP* de radio  $n+1$ , con  $n \geq 0$ , y el operador *LBP* de radio unitario aplicado a la escala  $1/2^n$ , ha sido contrastada mediante los resultados obtenidos. Concretamente, se verifica a través del experimento de la Sec. 5.3.1, en el que una utilización simultánea de múltiples niveles  $n$  de la **pirámide de imagen** proporciona mejoras relevantes en el poder predictivo respecto a la utilización del nivel  $n = 0$  de manera única, facilitando el análisis de vecindades de mayor dimensión. Así mismo, la posibilidad de utilizar la información generada por el operador *LBP* como herramienta principal de segmentación de imágenes ha sido confirmada mediante los resultados obtenidos a lo largo del trabajo, tanto en los experimentos intermedios como, concretamente, en la Sec. 5.4, donde se confirma la capacidad discriminatoria de las características generadas.
- Una segunda hipótesis, relativa a la posibilidad de utilización de un **clasificador «tradicional»** para realizar la segmentación de imágenes a nivel de píxel por medio de la información de las características generadas por el operador

*LBP* se ha visto igualmente contrastada mediante los resultados obtenidos a lo largo de la experimentación y concretamente en la Sec. 5.4, en la que se encuentran los resultados de test en los diferentes datasets, alcanzando resultados competitivos con los de investigaciones previas no basadas en *Deep Learning*. Con ello, se logra establecer una metodología que proporcione un poder predictivo elevado para aquellas aplicaciones en las que no sea preciso alcanzar la mayor precisión posible y no se disponga de bases de datos de gran tamaño o de una *GPU* con la que llevar a cabo un costoso proceso de entrenamiento.

- De igual forma, se verifica el objetivo de independizar la metodología de las particularidades de las bases de datos mediante la **automatización** de los procesos en los que es preciso sintonizar una parametría propia a cada base de datos. Se logra definir una serie de procesos automáticos que adaptan la metodología a las propiedades de las diferentes bases de datos a través de técnicas de computación evolutiva para la etapa de preprocesamiento y generación de características en las Sec. 4.2.3.1 y 4.2.3.2, respectivamente, y, en el caso de la verificación de la mejor combinación de parámetros, por medio de la estrategia *Grid Search* (véase la Sec. 4.2.3.3).
- Dado el objetivo de **evaluar** los resultados en condiciones independientes de los procesos de experimentación, sintonización y entrenamiento en múltiples bases de imágenes, se han empleado datasets contrastados mediante una amplia literatura, con las mismas divisiones y técnicas de evaluación. Mediante la evaluación en imágenes independientes es posible verificar los resultados individuales de cada una de ellas, tanto en lo relativo a las métricas como a la visualización de las etiquetas asignadas a cada uno de los píxeles.
- Como última consideración relativa a los objetivos, se ha desarrollado una etapa de **postprocesamiento** que proporciona una segmentación con mayor continuidad, reduciendo el ruido existente. Ha sido realizada mediante la adición de una segunda fase de entrenamiento donde a partir de la primera iteración, en la que el tratamiento de los píxeles es independiente, se utilizan sus resultados predictivos para generar una segunda iteración, es decir, un «*boost*» del clasificador en el que la información de las predicciones en regiones vecinas es empleada para proporcionar una clasificación final «refinada» para cada píxel.

De manera adicional, de los resultados de los procesos de análisis, experimentación y evaluación de la segmentación es posible definir una serie de conclusiones secundarias:

- Resulta complejo establecer una comparación equitativa con respecto a la literatura dado que el número de investigaciones recientes basadas en clasificadores tradicionales con las mismas bases de datos es limitado. Donde algunas de ellas tampoco se encuentran evaluadas exactamente en las mismas condiciones. De la existencia de una literatura limitada se extrae la conclusión de que la metodología propuesta supone una investigación que abre una línea de investigación alternativa, que no ha sido estudiada en detalle previamente,

donde los estudios basados en *Deep Learning* predominan holgadamente entre las investigaciones más recientes.

- Los valores de intensidad más frecuentes en la imagen de partida, es decir, la existencia de una tonalidad más oscura o más clara, tienen un gran impacto sobre las prestaciones de la metodología desarrollada. Se trata de una conclusión extraída en las tres bases de datos evaluadas en la Sec. 5.4, siendo el elemento que más relevancia tiene al comparar los resultados entre diferentes retinografías de un mismo dataset, siendo un punto de mejora que será detallado en la próxima sección, relativa al trabajo futuro (véase la Sec. 6.2).
- Las patologías tienen un impacto relevante sobre el poder predictivo dado que en múltiples ocasiones generan una elevada cantidad de ruido que puede ser observada mediante resultados visuales.
- En lo relativo al operador *LBP*, cabe destacar que en su variante *VAR* aporta información complementaria a la de las restantes variantes del operador debido a su medición de varianza de intensidad. Aporta un incremento notable de las prestaciones del clasificador, como puede ser comprobado en la Tabla 5.5, donde el descarte de las características generadas mediante dicha variante del operador supone una pérdida muy relevante del poder predictivo. Así mismo, como resulta intuitivo, en el problema de aplicación de segmentación de vasos sanguíneos la orientación de estos no es relevante y por ello la información proporcionada por los operadores invariantes a la rotación es más relevante.

## 6.2. Trabajo futuro

Expuestas las conclusiones del trabajo y los resultados de este, se indican a continuación las principales líneas de trabajo futuro que han sido identificadas para dar continuación a la investigación realizada, con el objetivo de complementarla con nuevas técnicas que puedan incrementar el comportamiento predictivo del algoritmo de segmentación finalmente desarrollado.

- Dado que uno de los principales problemas identificados es la varianza en el nivel de intensidad medio de las imágenes, generando una gran disparidad de los resultados entre algunas de las retinografías de una misma base de datos, se considera de interés el análisis de metodologías que puedan solventar dicho inconveniente. Como posible solución, se propone la aplicación de una técnica similar a la del operador *CLAHE*, es decir, como fase final de la etapa de preprocesamiento llevar a cabo una de normalización del contraste basada en los histogramas de intensidades, adaptándolos entre el mismo rango de valores en todas las imágenes.
- Así mismo, dado que otro elemento que afecta a los resultados predictivos es la presencia de patologías, se propone la realización de un análisis de las formas

geométricas que tienen las predicciones realizadas para descartar aquellas que tengan baja probabilidad de constituir vasos sanguíneos. Patologías como las presentes en la Fig. 5.30 poseen una geometría diferente a la de los vasos sanguíneos, y, por lo tanto, mediante operaciones de postprocesamiento podrían verse mejoradas. Sin embargo, esta técnica únicamente sería de utilidad para el problema de aplicación, es decir, la segmentación de imágenes de retina.

- Estudio de operaciones de preprocesamiento adicionales o alternativas a las empleadas en este trabajo con la finalidad de que su proceso de sintonización optimizado de lugar a imágenes de partida desde las que el poder discriminatorio de las características generadas por medio del operador *LBP* se vea incrementado. Una mayor optimización de dicha fase puede ser relevante de cara a adaptar la metodología a nuevas bases de datos, al igual que las prestaciones del problema de aplicación.
- Análisis de la capacidad de utilización de la metodología desarrollada en otros problemas diferentes a la segmentación de vasos sanguíneos en retinografías, empleando bases de datos de otro tipo, incluso considerando problemas en los que haya que segmentar más de dos elementos, es decir, con un etiquetado multiclase de cada uno de los píxeles.

# A. Apéndice

## A.1. Resultados gráficos del proceso de constitución de la metodología

A continuación se encuentran los resultados gráficos de la aplicación de las diferentes herramientas definidas en el apartado 4.2, correspondiente a la definición de la metodología de segmentación de los vasos sanguíneos en las retinografías.

Dado que el proceso se ha realizado sobre la base de datos *DRIVE*, la cual se ha dividido en diferentes sets de datos, entre los que se encuentran entrenamiento, validación y test, y puesto que las imágenes que se muestran a continuación corresponden a las verificaciones de las hipótesis establecidas en la generación de la metodología, se proporcionan los resultados gráficos sobre una misma imagen correspondiente al conjunto de datos de validación.



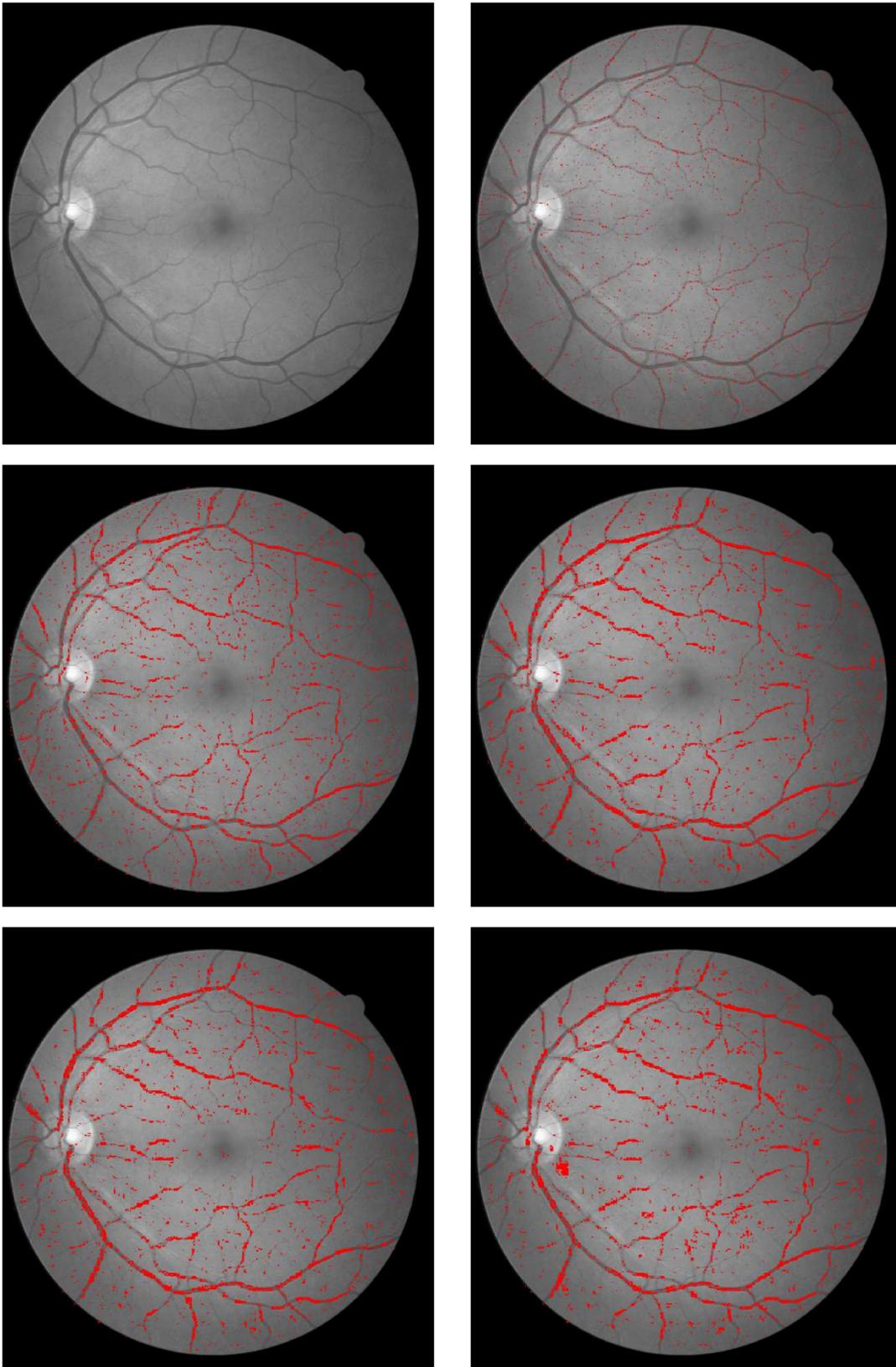
**Figure A.1.:** Imagen «35\_training.tif» del dataset *DRIVE* de manera previa a la aplicación del preprocesamiento.

Concretamente la imagen empleada es la denominada como «*35\_training.tif*», figura A.1, sobre la cual se aplican predicciones sobre los clasificadores entrenados en los procesos intermedios definidos, simulando la aplicación real de la herramienta de segmentación generada y facilitando de dicha forma la comparación tanto de la aplicación de cada uno de los procesos con diferente número de características como los propios procesos entre sí.

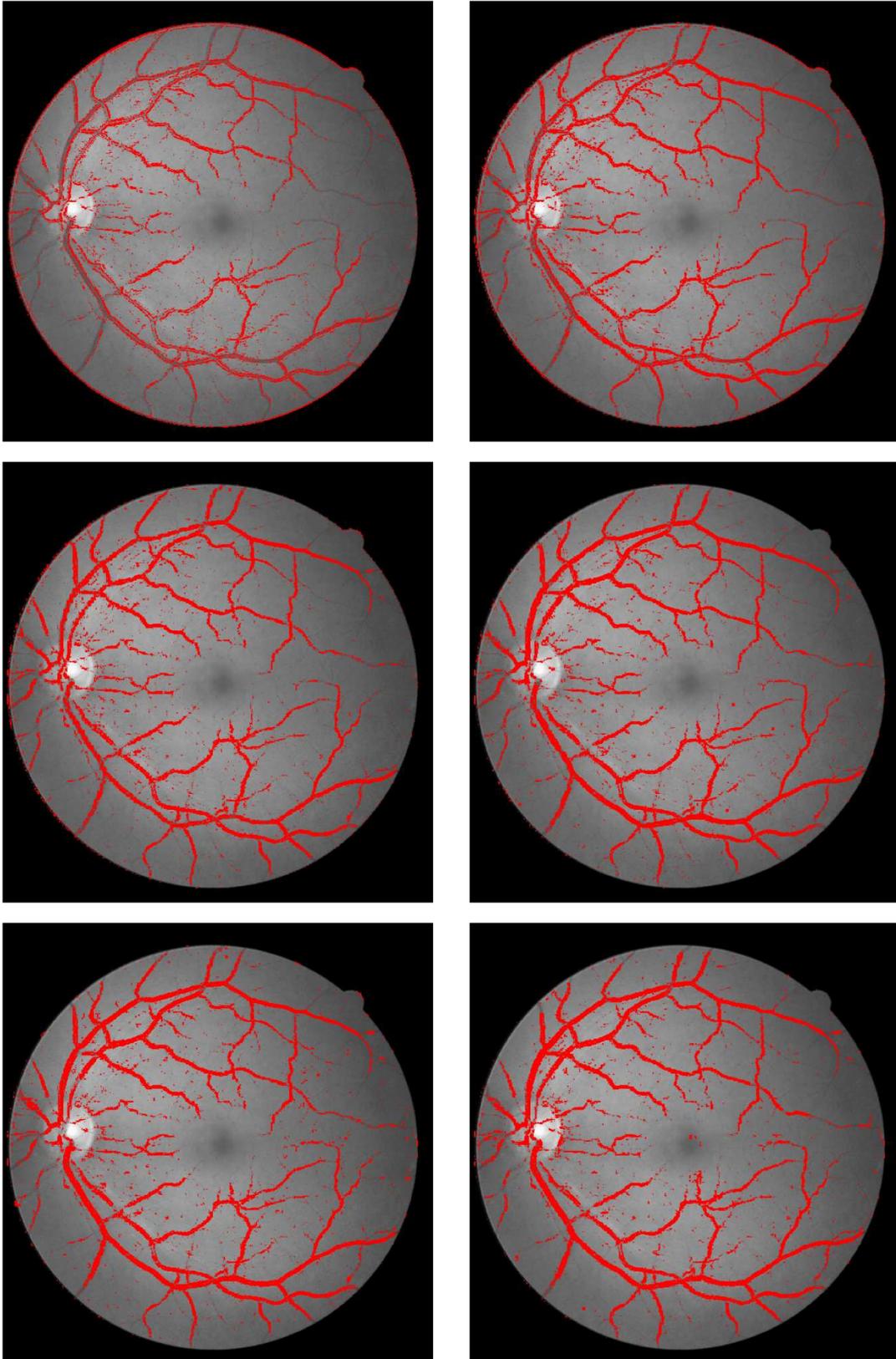
Dicha imagen puede ser encontrada en la figura A.2 tras la aplicación del correspondiente preprocesamiento manual que ha sido empleado a la definición de la metodología inicial, de manera previa a la aplicación de procesos optimizados mediante computación evolutiva.



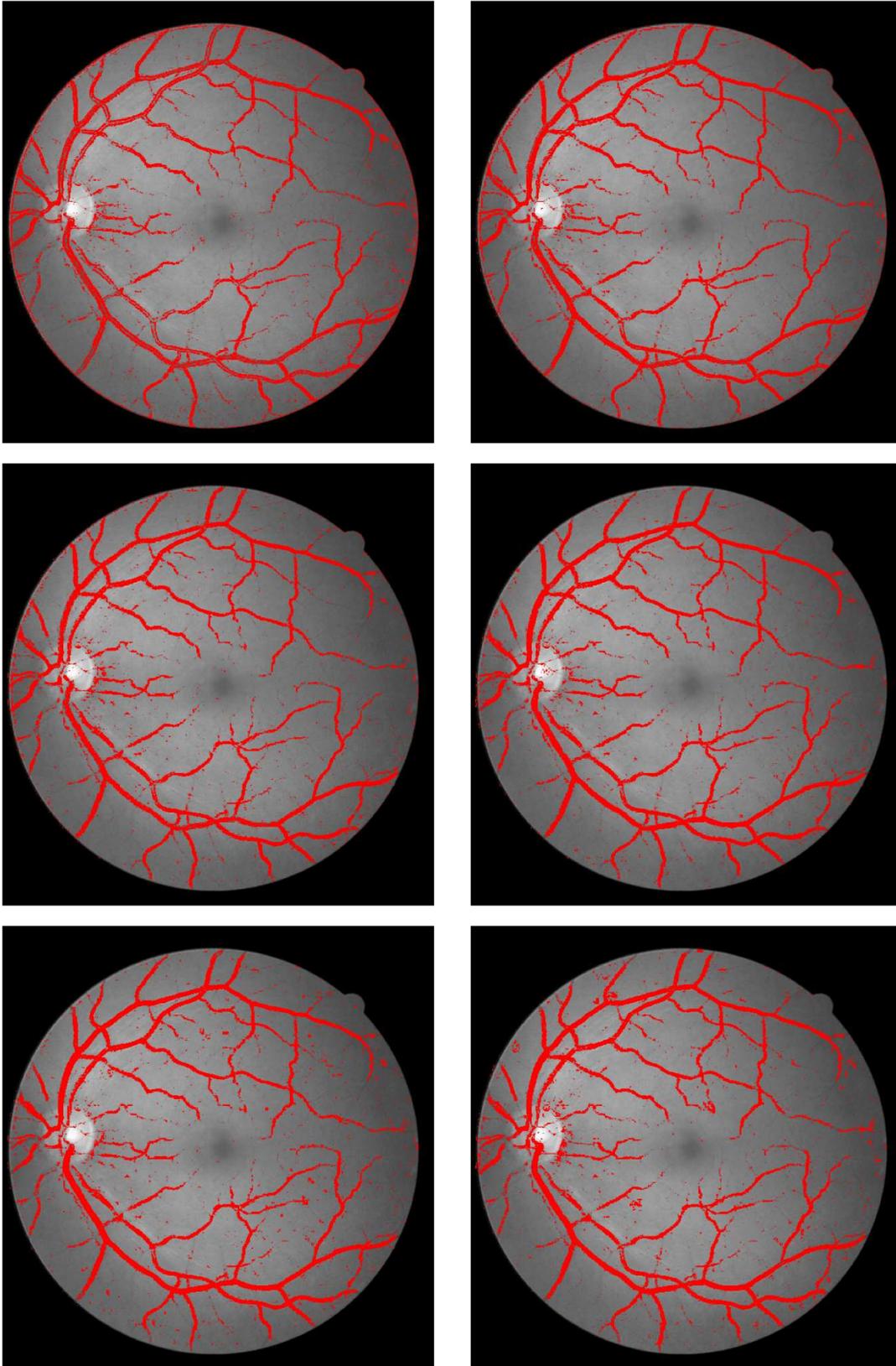
**Figure A.2.:** Imagen «*35\_training.tif*» del dataset *DRIVE* tras la aplicación del preprocesamiento manual.



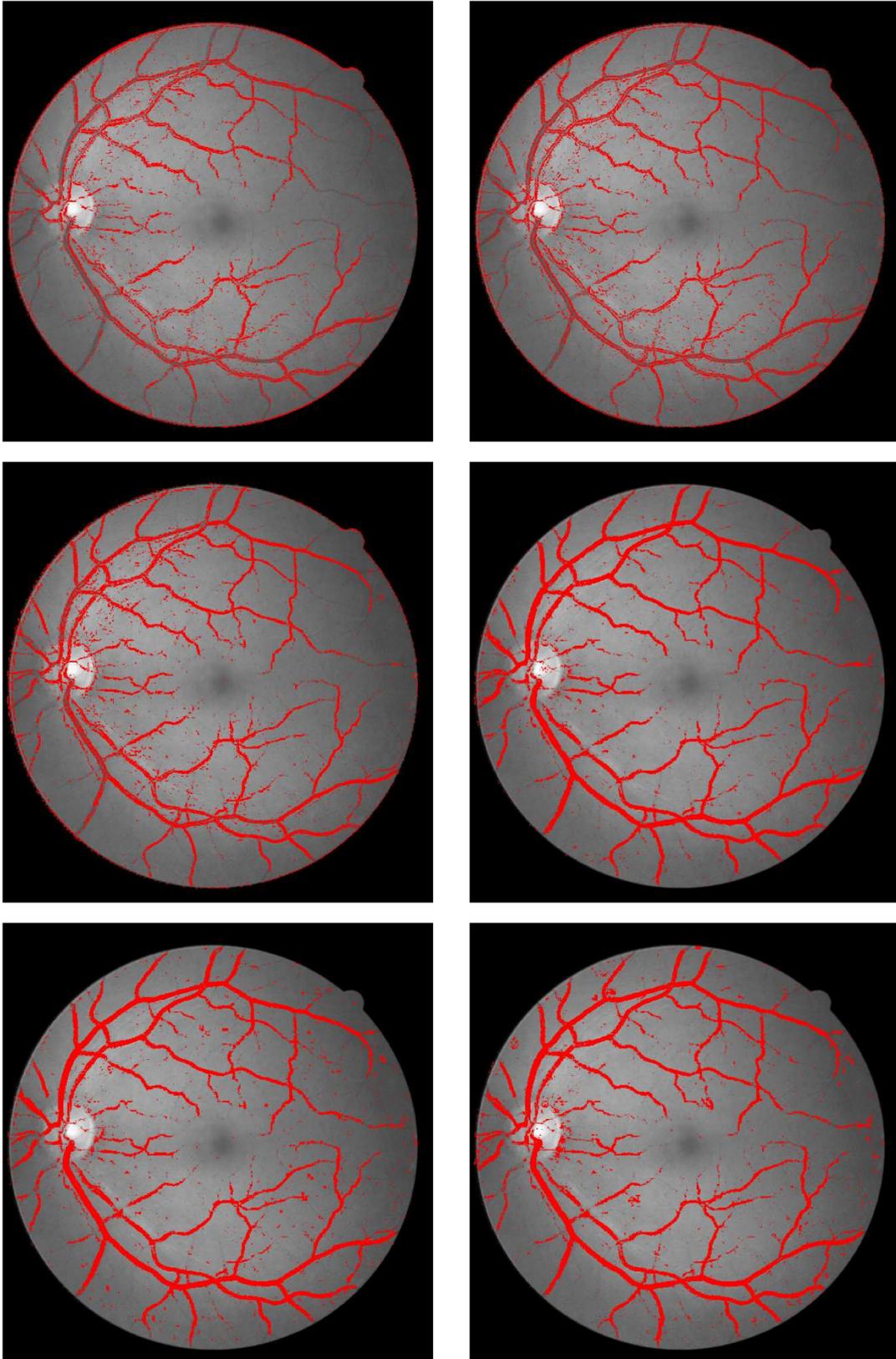
**Figura A.3.:** Representación de la predicción de los vasos sanguíneos (píxeles rojos) al aplicar el operador *LBP* a diferentes escalas en una imagen de validación de *DRIVE*. Resultados del experimento de la Sec. 5.3.1.



**Figura A.4.:** Representación de la predicción de los vasos sanguíneos (píxeles rojos) con las características de las diferentes variantes del operador *LBP* a múltiples escalas en una imagen de validación de *DRIVE*. Resultados del experimento de la 98Sec. 5.3.2.



**Figura A.5.:** Representación de la predicción de los vasos sanguíneos (píxeles rojos) con la inclusión de la característica correspondiente a la intensidad en una imagen de validación de *DRIVE*. Resultados del experimento de la Sec. 5.3.3.



**Figura A.6.:** Representación de la predicción de los vasos sanguíneos (píxeles rojos) con la inclusión de las características de la resolución 2 : 1 en una imagen de validación de *DRIVE*. Resultados del experimento de la Sec. 5.3.4.

# Referencias

- [1] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. *Lecture Notes in Computer Science*, 06 2000.
- [2] Sumit Saha. Xgboost vs lightgbm: How are they different. *Neptune AI*, July 2022.
- [3] Arden Dertat. Applied deep learning - part 4: Convolutional neural networks. *Towards Data Science*, November 2017.
- [4] Retinal vascular disorders. <https://www.singhealth.com.sg/patient-care/conditions-treatments/retinal-vascular-disorders>, July 2022.
- [5] Daniel Maturana, Domingo Mery, and Alvaro Soto. Face Recognition with Local Binary Patterns, Spatial Pyramid Histograms and Naive Bayes Nearest Neighbor Classification. *Paper presented at the international conference of the Chilean Computer Science Technology*, November 2009.
- [6] Yujia Jiang and Xin Liu. Uniform Local Binary Pattern for Fingerprint Liveness Detection in the Gaussian Pyramid. *Journal of Electrical and Computer Engineering*, 01 2018.
- [7] Martin Cerman, Ines Janusch, Rocio Gonzalez-Díaz, and Walter Kropatsch. Topology-based image segmentation using LBP pyramids. *Machine Vision and Applications*, 07 2016.
- [8] Arvind Yadav, R.s Anand, Mohan Dewal, and Sangeeta Gupta. Gaussian image pyramid based texture features for classification of microscopic images of hardwood species. *Optik - International Journal for Light and Electron Optics*, 09 2015.
- [9] Yimo Guo, Guoying Zhao, and Matti Pietikäinen. Texture classification using a linear configuration model based descriptor. 01 2011.
- [10] Esa Rahtu, Janne Heikkilä, Ville Ojansivu, and Timo Ahonen. Local phase quantization for blur-insensitive image analysis. *Image and Vision Computing*, 08 2012.
- [11] Yong-chun Miao and Yan Cheng. Automatic extraction of retinal blood vessel based on matched filtering and local entropy thresholding. *8th International Conference on Biomedical Engineering and Informatics*, October 2015.

- 
- [12] Zhenlin Guo, Ping Lin, Guangrong Ji, and Yangfan Wang. Retinal vessel segmentation using a finite element based binary level set method. *Inverse Problems and Imaging*, May 2014.
- [13] Yun Jiang, Huixia Yao, Zeqi Ma, and Jingyao Zhang. Bi-sanet - bilateral network with scale attention for retinal vessel segmentation. *Symmetry*, 09 2021.
- [14] Jingfei Hu, Hua Wang, Jie Wang, Yunqi Wang, Fang He, and Jicong Zhang. Sa-net: A scale-attention network for medical image segmentation. *PloS one*, 04 2021.
- [15] Jiang Yun, Hai Zhang, Ning Tan, and Li Chen. Automatic retinal blood vessel segmentation based on fully convolutional neural networks. *Symmetry*, 09 2019.
- [16] Shahab Aslani and Haldun Sarnel. A new supervised retinal vessel segmentation method based on robust hybrid features. *Biomedical Signal Processing and Control*, 06 2016.
- [17] Chengzhang Zhu, Beiji Zou, Zhao Rongchang, Jinkai Cui, Xuanchu Duan, Zailiang Chen, and Yixiong Liang. Retinal vessel segmentation in colour fundus images using extreme learning machine. *Computerized Medical Imaging and Graphics*, 05 2016.
- [18] Md. Abu Sayed, Sajib Saha, G. Rahaman, Tanmai Ghosh, and Yogesan Kanagasam. An innovate approach for retinal blood vessel segmentation using mixture of supervised and unsupervised methods. *IET Image Processing*, 11 2020.
- [19] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 01 1996.
- [20] Non rotation invariant local binary patterns. <https://github.com/scikit-image/scikit-image/>, July 2022.
- [21] Non gray scale invariant local binary patterns. <https://github.com/scikit-image/scikit-image/>, July 2022.
- [22] Vishal Morde. Xgboost algorithm: Long may she reign! *Towards Data Science*, April 2019.
- [23] Lidia Martínez. Comparativa de modelos de machine learning para la estimación de parámetros de interés empleando datos de la european soil database. Master's thesis, Universidad de Valladolid, September 2021.
- [24] Ricardo Recarey Fernández. Métodos de ensamblado en machine learning. Master's thesis, Universidad de Coruña, September 2021.
- [25] Cory Maklin. Gradient boosting decision tree algorithm explained. *medium*. *Towards Data Science*, May 2019.
- [26] Joaquín Amat Rodrigo. Regularización ridge, lasso y elastic net con python. *Ciencia de datos*, November 2020.

- [27] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, April 2004.
- [28] Drive: Digital retinal images for vessel extraction. <https://drive.grand-challenge.org/>.
- [29] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, Sep. 2012.
- [30] Retinal image analysis. <https://blogs.kingston.ac.uk/retinal/chasedb1/>.
- [31] A.D. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, March 2000.
- [32] Structured analysis of the retina. <https://cecas.clemson.edu/~ahoover/stare/>, September 2022.
- [33] Applying bilateral filter. <http://opencvexamples.blogspot.com/2013/10/applying-bilateral-filter.html>, June 2022.
- [34] Histograms - image processing. [https://docs.opencv.org/4.x/d6/dc7/group\\_\\_\\_imgproc\\_\\_\\_hist.html](https://docs.opencv.org/4.x/d6/dc7/group___imgproc___hist.html), June 2022.
- [35] Resize filters. <https://pillow.readthedocs.io/en/stable/handbook/concepts.html>, May 2022.
- [36] Vinny DaSilva. Computer vision for busy developers. *Medium*, August 2019.
- [37] ArIES. Texture analysis using LBP. *Medium*, August 2018.
- [38] Roberto Canduela. Creación de variables sintéticas mediante evolución gramatical en problemas de clasificación. Master's thesis, Universidad Nacional de Educación a Distancia, October 2017.
- [39] Federico Castejón. Síntesis automática de amplificadores electrónicos mediante Grammatical Evolution. Master's thesis, Universidad Nacional de Educación a Distancia, June 2013.
- [40] Antonio Rodríguez. Segmentación automática de lared de vasos sanguíneos en imágenes de retina mediante redes neuronales artificiales evolutivas y operadores basados en local binary patterns (LBP). Master's thesis, Universidad Nacional de Educación a Distancia, September 2013.
- [41] A. Eiben and Jim Smith. *Introduction To Evolutionary Computing*. 01 2003.