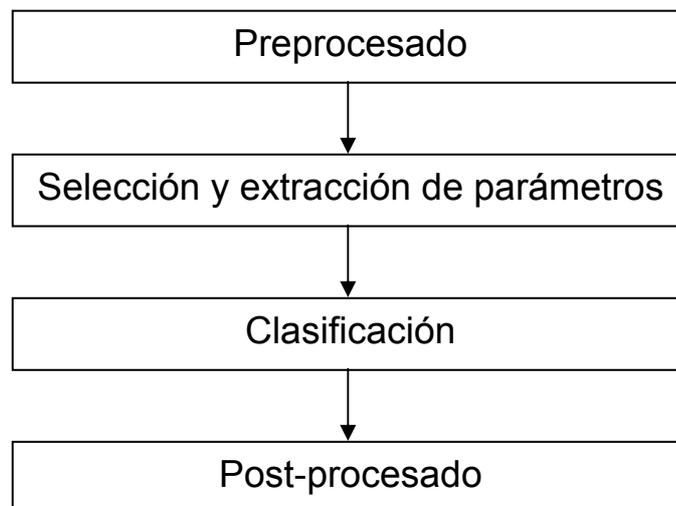


CAPÍTULO 4. Descripción del algoritmo

4.1. Introducción

El algoritmo objeto de este proyecto consta de 4 fases, tal y como se muestra en el esquema 4.1. En la primera etapa o preprocesado, se lleva a cabo una mejora de la imagen para obtener una normalización adecuada en la iluminación y el contraste. En el siguiente paso, se extraen grupos de entrenamiento dinámicos propios de cada imagen. Posteriormente, los píxeles se clasifican en dos clases usando un Discriminante Lineal de Fisher (FLD, del inglés *Fisher's Linear Discriminant*). Y por último, el resultado se post-procesa para eliminar los falsos positivos, tales como las manchas blancas algodonosas, las drusas, el disco óptico, etc.



Esquema 4.1. Diagrama de bloques

4.2. Etapa de preprocesado

Las imágenes de retina presentan una alta correlación entre la pigmentación de la piel y el color del iris, por lo que el color del fondo es típicamente único para cada individuo. Además estas imágenes son pobres en

contraste y la iluminación desigual provoca niveles de intensidad mayores en las regiones cercanas al disco óptico y menores en las regiones más alejadas. Todas estas características tienen un impacto significativo en la detección de lesiones. Por tanto, es necesario normalizar el color y mejorar el contraste antes de empezar con la detección de exudados duros. En este caso, se usa una modificación del modelo RGB con el fin de obtener estas mejoras. Con él se evitan los problemas asociados al aplicar métodos en escala de grises a cada una de las componentes debido a la alta correlación entre ellas. La transformación propuesta es la siguiente:

$$\begin{aligned}
 (R, G, B) &\rightarrow (Y, I, Q) \\
 Y_{mod} &= aY + bI + cQ \\
 (Y_{mod}, I, Q) &\rightarrow (R_{mod}, G_{mod}, B_{mod})
 \end{aligned} \tag{1}$$

donde (R,G,B) son las tres componentes del modelo RGB original, (Y,I,Q) las componentes del modelo YIQ original y $(R_{mod},G_{mod},B_{mod})$ las componentes del modelo modificado RGB, RGB_{mod} . Las constantes a , b y c son parámetros que dependen de las características de las imágenes. Se seleccionan para reducir la desviación estándar local de la luminancia (σ_μ) en Y_{mod} y para aumentar las medias (μ_σ y μ_c) del contraste [10]. Para nuestro caso, tomando unos valores para $[a,b,c]=[1.5,-1,1]$, se obtiene una mejora significativa en la saturación total del color y en el contraste entre las lesiones y el fondo de la imagen, tal y como se puede observar en las figuras 4.1 y 4.2.



Figura 4.1. Imagen original



Figura 4.2. Imagen obtenida tras el preprocesado

El motivo por el que se elige el modelo RGB_{mod} y no otro de todas las posibilidades existentes, es que con él se obtiene experimentalmente el mejor valor para la métrica cuantitativa J .

$$J = \text{traza} \left(\frac{S_b}{S_w} \right), \quad (2)$$

donde S_b es la matriz de covarianza entre los miembros de las diferentes clases y S_w la matriz covarianza dentro de los miembros de una misma clase, cuyas expresiones se verán en detalle en el Apartado 4.4.

Con esta métrica se estima la separación de clases entre los píxeles exudados y los píxeles no exudados. Un valor alto de J indica que las dos clases (exudados y no exudados) están más separadas mientras que los miembros de una misma clase están más cercanos entre ellos. En la tabla 4.1, se muestran los valores obtenidos para varios modelos. Como se puede comprobar, para el modelo aquí definido se consigue el valor más alto.

<i>Modelo de color</i>	<i>RGB</i>	<i>HSI</i>	<i>YIQ</i>	<i>YC_bC_r</i>	<i>Lab</i>	<i>Luv</i>	<i>RGB_{mod}</i>
<i>J</i>	2.91	2.98	2.91	2.92	2.99	2.49	4.14

Tabla 4.1. Métrica separabilidad entre clases, J

4.3. Selección y extracción de características

Para poder clasificar cada píxel y decir si se trata o no de un exudado, se necesita conocer a priori las características que presentan los píxeles exudados. Como los exudados duros se caracterizan principalmente por su color, el espacio característico seleccionado se define a partir de las componentes del modelo descrito en la sección anterior. Por tanto, para cada píxel, se definirá un vector característico \mathbf{X} compuesto por las tres componentes del modelo RGB_{mod} , ya que es en este espacio donde más separadas están las dos clases existentes:

$$\mathbf{X} = [R_{mod}, G_{mod}, B_{mod}] \quad (3)$$

En una clasificación basada en dos clases, hay que conocer este vector característico para los píxeles exudados y para los no exudados. Es decir, hay

que enseñarle a distinguir entre las dos clases posibles para que el clasificador aprenda por sí mismo a agrupar los píxeles. Normalmente, esto se hace pasando al clasificador dos secuencias de entrenamiento, una por clase, que se obtienen de un grupo de imágenes de prueba en las que se les indica qué es y qué no es un exudado. Debido a la gran variabilidad intraclases entre imágenes, incluso después de la etapa de preprocesado, las secuencias de datos así obtenidas no son representativas de las características de todas las imágenes, y por tanto, el clasificador no aprende. En este trabajo se propone extraer los grupos de entrenamiento automáticamente para cada imagen, evitando la segmentación manual. Entonces, para cada imagen el algoritmo escoge los respectivos grupos de entrenamiento, exudado y no exudado.

El grupo de entrenamiento exudado se compone por los píxeles que pertenecen a los pequeños exudados aislados. Éstos presentan las mismas características que el resto de exudados de la imagen y además se pueden encontrar en todos los fondos de imágenes con exudados duros. El grupo de entrenamiento exudado se obtiene realizando una segmentación gruesa de la imagen. Primero, se realzan los bordes del exudado aplicando el operador de Frei-Chen [11] a la componente R_{mod} , que destaca los bordes independientemente de la magnitud del brillo:

$$I_{frei} = \cos^{-1} \left[\frac{\sum_{i=1}^4 (v_i \cdot r_{mn})^2}{\sum_{i=1}^9 (v_i \cdot r_{mn})^2} \right]^{1/2} \quad (4)$$

donde r_{mn} es un vector que representa una subimagen 3x3 de R_{mod} centrado en el píxel (m,n); v_i con $i=1..9$ son las máscaras de Frei-Chen en \mathfrak{R}^9 y v_i con $i=1..4$ forma el subespacio de bordes de Frei-Chen. A continuación, se muestran las máscaras de Frei-Chen:

$$\begin{aligned}
 v_1 &= \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix} & v_2 &= \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix} & v_3 &= \begin{bmatrix} 0 & -1 & \sqrt{2} \\ 1 & 0 & -1 \\ -\sqrt{2} & 1 & 0 \end{bmatrix} \\
 v_4 &= \begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -\sqrt{2} \end{bmatrix} & v_5 &= \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} & v_6 &= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \\
 v_7 &= \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} & v_8 &= \begin{bmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{bmatrix} & v_9 &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}
 \end{aligned}$$

El subespacio formado por v_i con $i=1\dots4$ se utiliza para detectar la presencia de bordes., denominándose subespacio de bordes. El subespacio formado por v_i con $i=5\dots8$ se utiliza para detectar la presencia de líneas y, por tanto, se conoce como subespacio de líneas. Y por último, v_9 es el operador de la media. Para el problema en cuestión, el subespacio de interés es el primero de ellos.

El operador de Frei-Chen proporciona un suavizado además del efecto de derivación. Ya que la derivación acentúa el ruido, el efecto de suavizado es particularmente interesante puesto que elimina parte del ruido.

Lo aplicamos a R_{mod} porque los exudados aparecen con una intensidad más alta y con un fondo más homogéneo.

En la figura 4.3 se muestra el resultado obtenido. Los exudados aparecen como una región blanca rodeada por un área de gris oscura, especialmente los pequeños exudados aislados (véase la figura 4.4). Para el correcto visualizado de la imagen I_{frei} se le realiza una ecualización del histograma, ya que sólo los bordes aparecen más oscuros, mientras que el resto de la imagen es demasiado clara.

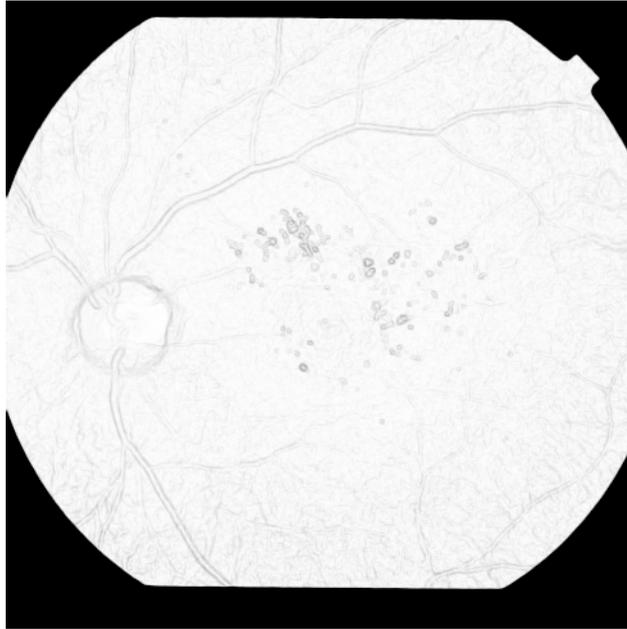


Figura 4.3. Resultado aplicar el operador Frei-Chen a R_{mod}

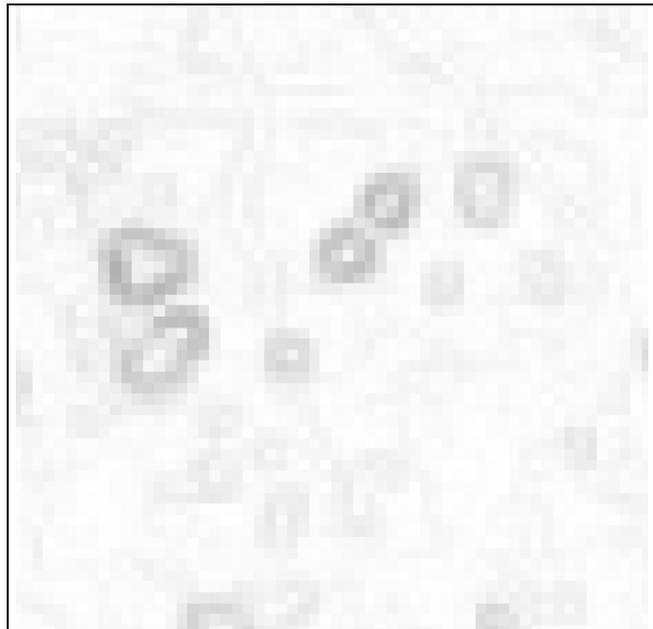


Figura 4.4. Zoom sobre I_{frei} : Ejemplo de exudados aislados

Tratando el resultado como un mapa de altura, los exudados aislados se pueden identificar como los grupos de píxeles brillantes rodeados por curvas cerradas (isocurvas) de descenso. Estos píxeles, $\{X_{ex}^1, \dots, X_{ex}^N\}$, los cuales se muestran en la figura 4.5, forman el grupo exudado de entrenamiento.



Figura 4.5. Secuencia de entrenamiento de los píxeles exudados (en verde sobre la imagen original)

El tamaño N varía de una imagen a otra dependiendo de la segmentación y del número de píxeles aislados encontrados.

Para los píxeles no exudados se forma el grupo de entrenamiento no exudado eligiendo M píxeles alrededor del disco óptico, donde M depende del área del disco óptico. Aunque estos píxeles no son fuertemente representativos de la clase no exudado, están muy cercanos a los píxeles exudados en el espacio característico [1]. Por lo tanto, el límite de decisión se desplaza más exactamente entre las dos clases, obteniendo un mejor funcionamiento en la tarea de clasificación.

Para localizar el disco óptico se tendrá en cuenta que se trata de una región brillante y donde entran los vasos sanguíneos y los nervios ópticos de la retina de forma vertical. El método para detectar el disco óptico se compone de tres partes: una primera donde se seleccionan los candidatos a disco óptico mediante morfología matemática; en la segunda se detecta el centro del disco óptico usando la transformada de Hough; y por último, se utiliza *active contour* para obtener el contorno del mismo. Este último paso ha sido desarrollado por

investigadores del Departamento de Teoría de la Señal y Comunicaciones de la Escuela Superior de Ingenieros de la Universidad de Sevilla.

En primer lugar, se aplican filtros secuenciales alternativos (*ASF*, de *Alternating Sequential Filter*) a la componente de la luminancia modificada, Y_{mod} , para calcular la aproximación del fondo:

$$Y_{ASF} = ASF(Y_{mod}) = \phi^{(nE)}(\dots(\gamma^{(2E)}(\phi^E(\gamma^E(Y_{mod}))))\dots) \quad (16)$$

donde ϕ y γ denotan el cierre y la apertura de imágenes grises respectivamente, y nE es un elemento estructurante del tamaño n suficientemente grande para quitar los elementos brillantes pero no el disco óptico. En la práctica, teniendo en cuenta el tamaño de las imágenes y del disco óptico, se ha observado un buen resultado con un valor de n de 20 iteraciones.

Como resultado se obtiene una imagen borrosa, tal y como la figura 4.6.

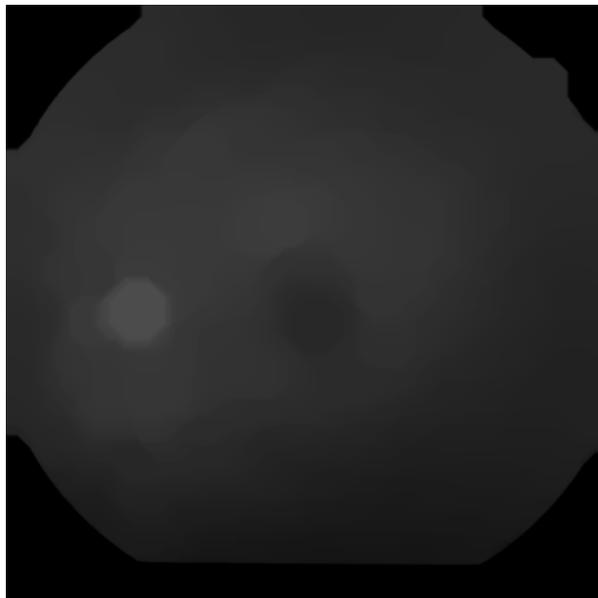


Figura 4.6. Imagen filtrada con filtros secuencias alternativos

Calculando los máximos locales de esta imagen resultante, se obtienen los centroides de las regiones candidatas a disco óptico. Como la imagen es

relativamente grande, al calcular los máximos locales se encuentra un número elevado de puntos que, al estar muy próximos entre ellos, se agrupan en tres o cuatro candidatos aplicando un algoritmo de distancias *K-means* (ver figura 4.7). De esta forma, queda un pequeño número de candidatos.

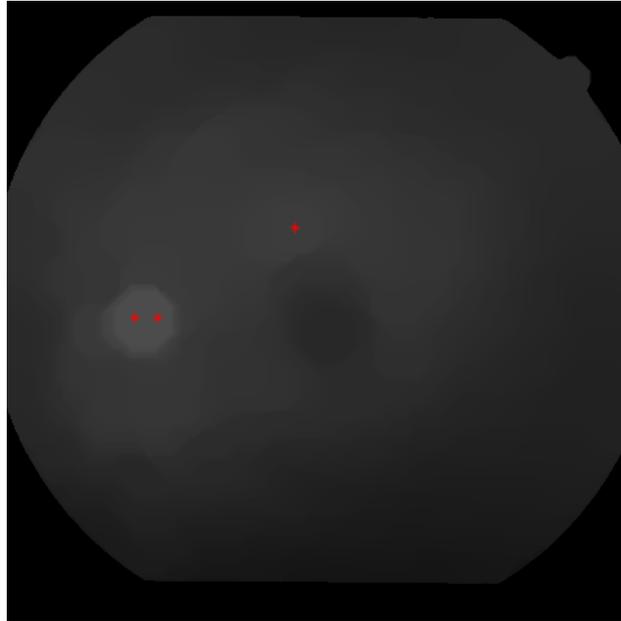


Figura 4.7. Representación candidatos a disco óptico

Otra característica a tener en cuenta para localizar el disco óptico, según se ha dicho anteriormente, es la presencia de vasos verticales en su vecindad. Se aplica entonces un filtro bidimensional orientado verticalmente caracterizado por una sección de perfil cruzado, tal y como se describe en [14]. Aquí se modela el vaso sanguíneo como una curva *gaussiana*.

$$K_i(x, y) = -\exp\left(-\frac{u^2}{2\sigma^2}\right) - m_i, \quad \forall p_i \in N \quad (17)$$

donde $N = \{(u, v), |u| \leq 3\sigma, |v| \leq L/2\}$ y $m_i = \frac{\sum_{p_i \in N} K_i(x, y)}{A}$, donde A es el número de puntos pertenecientes a N .

Tomando:

$$\begin{aligned}\sigma &= 2 \\ L &= 9 \\ \theta_i &= \frac{\pi}{2}\end{aligned}\tag{18}$$

se obtienen diferentes filtros para las distintas orientaciones. En particular, el filtro en la dirección vertical es el siguiente:

$$k_{\frac{\pi}{2}} = \begin{bmatrix} 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & -2 & -5 & -6 & -5 & -2 & 1 & 2 & 3 & 4 \end{bmatrix}\tag{19}$$

De esta forma, se detectan los vasos verticales que hay en la imagen, en este caso, en Y_{mod} , tal y como se muestra en la figura 4.8.



Figura 4.8. Líneas verticales detectadas

A continuación, se usa la Transformada de Hough [12] para modelar cada vaso que ha quedado en la imagen filtrada como una línea simple para así obtener el esqueleto de la misma. Esta transformada consiste en una herramienta para encontrar la ecuación de una línea que pase por un conjunto de m puntos en el plano (ρ, θ) . Una línea que une una secuencia de puntos o píxeles se puede expresar de la siguiente forma:

$$\rho = x \cdot \cos \theta + y \cdot \text{sen} \theta \quad (20)$$

El objetivo es encontrar los parámetros ρ y θ . Para ello, es necesario discretizar el espacio de parámetros en una serie de celdas denominadas celdas de acumulación. Esta discretización se realiza sobre los intervalos $(\rho_{\min}, \rho_{\max})$ y $(\theta_{\min}, \theta_{\max})$ (véase la figura 4.9).

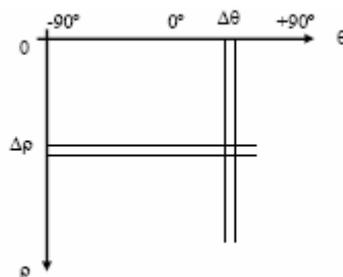


Figura 4.9. Discretización plano (ρ, θ)

El siguiente paso es evaluar la ecuación de la recta para cada punto de la imagen, (x_k, y_k) . Si se cumple la ecuación, se incrementa en uno el número de votos en el acumulador de la celda correspondiente. A final del proceso, un número elevado en el acumulador indica que el punto pertenece a la recta. Se usan las coordenadas polares y no las cartesianas para evitar problemas asociados a estas últimas: en coordenadas cartesianas, la ecuación de una recta es:

$$y = a \cdot x + b \quad (21)$$

En este caso, se discretiza el espacio de parámetros (a, b) . Se presenta un problema importante a la hora de elegir los límites (a_{\min}, a_{\max}) , ya que a priori no

son conocidos. Además, tanto la pendiente (a) como la ordenada (b) se acercan al infinito cuando la recta se aproxima a posiciones verticales. Usando las coordenadas polares se consigue barrer todo el rango de variación para θ desde -180° a 180° .

De esta forma, el resultado de aplicar la Transformada de Hough a la imagen filtrada verticalmente contiene las líneas detectadas, tal y como se puede ver en la figura 4.10.

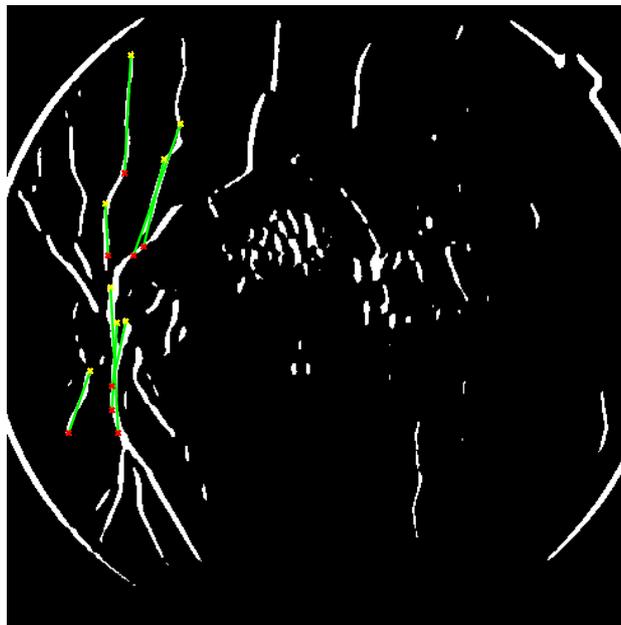


Figura 4.10. Representación líneas detectadas con la Transformada de Hough sobre las líneas verticales detectadas con Kirsch

De todos los puntos candidatos a disco óptico obtenidos con la morfología matemática, se escogerá como centroide del disco óptico al que tenga mayor número de líneas que pasen por él (Figura 4.10).

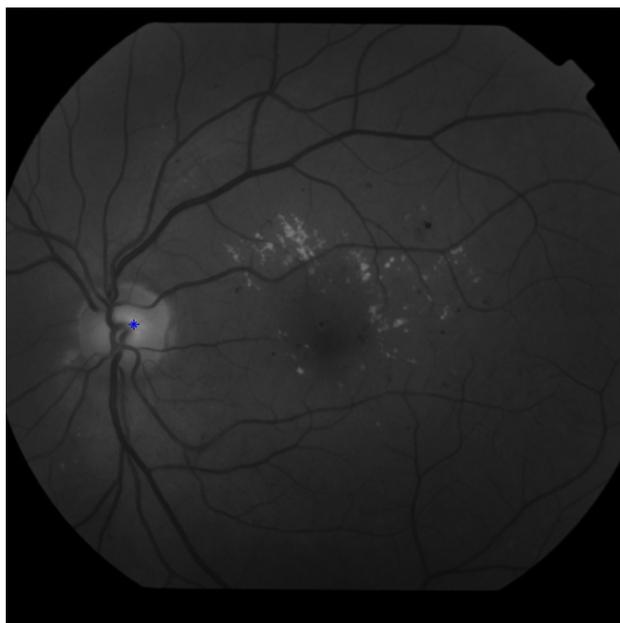


Figura 4.10. Localización disco óptico

Es muy importante localizar y detectar el contorno del disco óptico para eliminar su efecto sobre el resultado del algoritmo. Todavía no se ha acabado con la detección del DO, ya que para quitar todos los puntos internos a él hay que conocer su extensión completa. En este sentido, investigadores de la Escuela Superior de Ingenieros de la Universidad de Sevilla, en concreto, F.J. Nuñez, I. Fondón, M. Tirado, C. Serrano, B. Acha, S. Jiménez y P. Alemany, han desarrollado un algoritmo pendiente de publicación, donde se incluye la detección del disco óptico mediante *active contours*. Una segmentación basada en esta técnica permite que un determinado contorno se modifique de forma que se minimice una función de energía. Entre los métodos existentes, se ha escogido uno basado en regiones, en los que se modela el objeto y el fondo de forma estadística y se encuentra el valor óptimo de energía, que es el que mejor modela la imagen. Para evitar los inconvenientes propios de utilizar modelos globales basados en región en imágenes con el fondo y primer plano con modelos estadísticos similares, donde la segmentación puede ser errónea, se utilizan modelos locales [15]. Se considera por separado cada punto y se minimiza la función de energía en su región local propia.

El número de iteraciones es 80. Al resultado obtenido, se le aplica una apertura (operación morfológica) para corregir algunos casos en los que aparecen unos

brazos erróneos sobresalientes del contorno del disco óptico. El elemento estructurante es un círculo de tamaño n , como en (16).

Tras el proceso completo de detección del disco óptico, se obtiene una matriz binaria, cuyos elementos a uno indican los puntos que pertenecen al interior del disco óptico, tal y como se muestra en la figura 4.10.

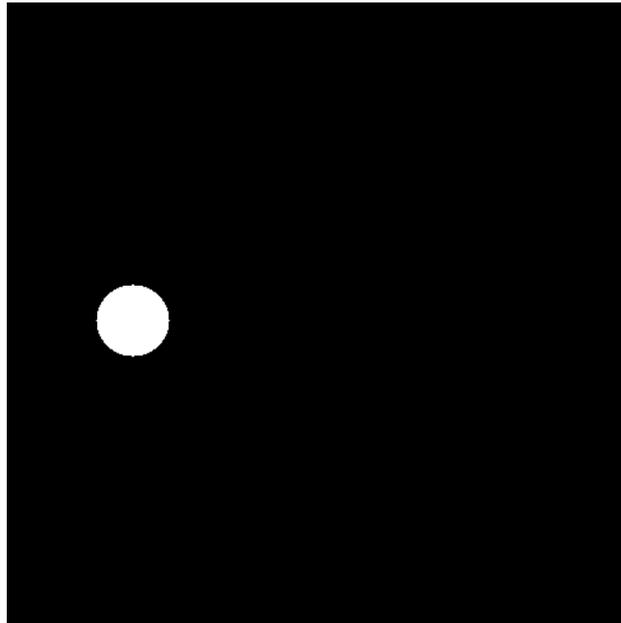


Figura 4.11. Detección del disco óptico completo usando *active contours*

La imagen obtenida como disco óptico se dilata levemente para evitar los puntos amarillentos alrededor del mismo o los puntos próximos al borde pero pertenecientes a él que se hayan eliminado tras la apertura aplicada. Esto se puede hacer sin problemas, ya que los exudados duros no suelen localizarse cerca del disco óptico.

4.4. Clasificación

Una vez se tienen los grupos de entrenamiento, tanto de los píxeles exudados como de los no exudados de la imagen en particular, se ha de proceder a clasificar todos los píxeles que forman la imagen. Para ello se

emplea un discriminante lineal de Fisher, debido a que éste tiene en cuenta la forma de los *clusters* o racimos en el espacio característico y a su eficiencia en el tiempo. Esto último es muy importante debido a que el algoritmo se aplicará sobre una cantidad considerable de imágenes, por lo que es aconsejable conseguir que no tarde demasiado en presentar los resultados.

El discriminante lineal de Fisher busca una dirección w de los datos en el espacio característico que maximice la separación entre las dos clases. Esta dirección w se encuentra maximizando una función S , que en cierta medida, no es sino una relación señal a ruido para el etiquetado de clases:

$$S(w) = \frac{w^T S_b w}{w^T S_w w} \quad (22)$$

donde S_b es la matriz de covarianza inter-clase, es decir, entre miembros de las dos clases diferentes que existen:

$$S_b = (\mu_{ex} - \mu_{no-ex})(\mu_{ex} - \mu_{no-ex})^T \quad (23)$$

y S_w la matriz de covarianza intra-clase o covarianza dentro de los miembros de una misma clase:

$$S_w = S_{ex} + S_{no-ex} \quad (24)$$

Siendo μ_{ex} y μ_{no-ex} las medias de las clases exudado y no exudado respectivamente, calculadas usando el correspondiente grupo de entrenamiento:

$$\mu_{ex} = \frac{1}{N} \sum_{i=1}^N X_{ex}^i \quad (25)$$

$$\mu_{no-ex} = \frac{1}{M} \sum_{i=1}^M X_{no-ex}^i \quad (26)$$

y las matrices de dispersión de cada clase:

$$S_{ex} = \sum_{i=1}^N (X_{ex}^i - \mu_{ex})(X_{ex}^i - \mu_{ex})^T \quad (27)$$

$$S_{no-ex} = \sum_{i=1}^M (X_{no-ex}^i - \mu_{no-ex})(X_{no-ex}^i - \mu_{no-ex})^T \quad (28)$$

De optimizar la función (5) resulta la siguiente regla lineal:

$$\delta_M^2 = (X - \mu_j)^T S_p^{-1} (X - \mu_j) \quad (29)$$

para $j = ex, no-ex$ y S_p , matriz de covarianzas ponderada, dada por la siguiente fórmula:

$$S_p = \frac{(N-1)S_{ex} + (M-1)S_{no-ex}}{N+M-2} \quad (30)$$

Cada píxel se clasificará como la clase respecto a la cual obtenga una distancia de Mahalanobis (12) menor. Este procedimiento es óptimo para clasificar entre dos poblaciones.

En la figura 4.12 se puede observar la imagen resultante de la tarea de clasificación, I_{class} , usando la regla anterior. La matriz que representa esta imagen es una matriz binaria, donde sus componentes o píxeles a uno indican que pertenecen a la clase exudado y los que están a cero indican que pertenecen a la clase no-exudado. A simple vista, a los píxeles más amarillentos se les asigna la clase exudado y a los más rojizos la clase no exudado.

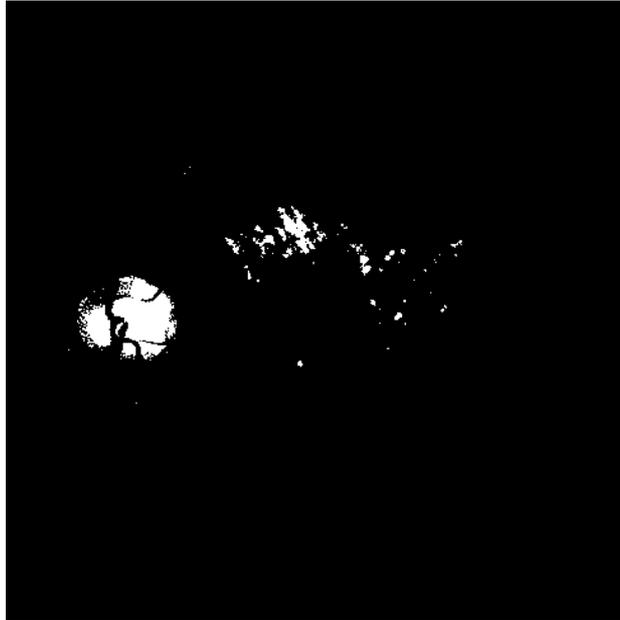


Figura 4.12. Imagen resultante de la clasificación

4.5. Post-procesado

No se puede dar por concluido el algoritmo, puesto que como se puede comprobar en la figura 4.12, después de la clasificación otras lesiones amarillentas, el disco óptico y zonas cercanas a él, se pueden detectar como EDs. En la etapa de post-procesado se pretende eliminar los falsos positivos detectados en las etapas previas y proporcionar así robustez al algoritmo frente a otros tipos de lesiones o manchas que están presentes en casi todas las imágenes de retina, como es el caso del disco óptico.

Anteriormente, se ha tenido en cuenta la principal característica de los exudados, como es su color. Pero otra cualidad significativa de los exudados es la agudeza de sus bordes; tienen unos bordes bien delimitados. En cambio, esta característica no la presentan otras lesiones como puntos algodonosos o lesiones en las regiones papilares. Por lo tanto, se puede hacer uso de esto durante el post-procesado para eliminar las regiones amarillentas detectadas en la fase de clasificación.

Primero, los bordes de la imagen se realzan usando esta vez el método de Kirsch a la componente verde original, G . El operador de Kirsch delinea los límites dándoles un valor dependiendo de la fuerza del borde, por lo que los exudados aparecerán más marcados que el resto de manchas presentes en la imagen.

$$I_{Kirsch}(m,n) = \max_i \{ (g_{mn} \cdot k_i) \} \quad (31)$$

donde g_{mn} es un vector que representa una subimagen 3x3 de la componente verde original centrada en el píxel (m,n) y k_i con $i=1..8$ son el grupo de máscaras que forman el operador de Kirsch [12]:

$$k_1 = \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad k_2 = \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad k_3 = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$k_4 = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad k_5 = \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad k_6 = \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$$

$$k_7 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad k_8 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

Para cada punto de la imagen se obtienen ocho valores resultantes de la convolución con cada una de las máscaras. El valor del módulo del gradiente resulta ser el máximo de esos ocho valores, mientras que la dirección queda determinada por el ángulo asociado a la máscara que ha generado dicho valor máximo, con la siguiente correspondencia:

Máscara	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
Ángulo (°)	0	45	90	135	180	225	270	315

Tabla 4.2. Ángulo asociado a cada máscara

Las máscaras de Kirsch se obtienen de una máscara simple rotándola en las ocho direcciones principales de la brújula: N, NO, O, S, SE; E, NE. De ahí que también se las llame máscaras de brújula.

En este algoritmo, sólo se tendrá en cuenta el módulo, por lo que se ignorará el ángulo asociado, ya que la orientación del borde no es relevante. Se ha escogido el operador de Kirsch porque, junto al método de Sobel, es el que produce menos errores en imágenes con ruido, comparándolo con otros modelos de gradiente.

Entonces, aplicando este operador a la componente verde, que es el plano donde más información hay y mejor es el contraste para la detección, el algoritmo delinea los bordes dándoles un valor dependiendo de la fuerza del mismo. Umbralizando el resultado a un nivel de gris α_1 se obtienen solamente los bordes más agudos, como se muestra en la figura 4.13:

$$I_{fuerza_borde} = T_{\alpha_1}(I_{Kirsch}) \quad (32)$$

α_1 es un parámetro del algoritmo que controla la compensación entre la sensibilidad y el número de falsos positivos. De nuevo, esta imagen o mejor dicho, la matriz que la representa, es binaria, cuyos valores a uno indican que se trata de un borde. Un valor demasiado alto en el umbral implicaría que el algoritmo no detecte todos los exudados duros de la imagen y un valor demasiado bajo detectaría como exudados otras lesiones que no lo son, ya que se estaría considerando un borde que no tiene tanta agudeza. Por lo tanto, en esta relación de compromiso, habrá que elegir correctamente un valor adecuado para α_1 , tal que, en la medida de lo posible, detecte todos los exudados presentes en la imagen pero no las demás anomalías.

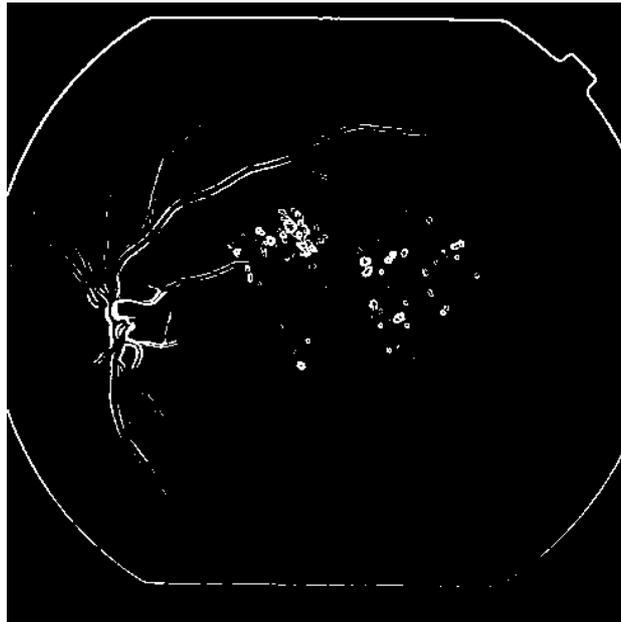


Figura 4.13. Detección de bordes más fuertes

En la práctica, se comprueba que el algoritmo alcanza el mejor funcionamiento usando un umbral de segmentación, α_1 , de 0.35.

Una vez se tienen identificadas las zonas con los bordes más acentuados, se utiliza esta información para que sólo se tengan en cuenta los píxeles clasificados como exudados que tengan un borde acentuado. Para ello, el resultado de la clasificación, I_{class} , y la imagen con los bordes, $I_{fuerza-borde}$, se combinan con reconstrucción por dilatación para eliminar los falsos positivos amarillos.

La reconstrucción por dilatación de una imagen máscara g por una imagen marcador f ($D_f = D_g$ y $f \leq g$) se define como la iteración de dilataciones geodésicas de f con respecto a g hasta la estabilidad, es decir, hasta que se obtenga la misma imagen, y se denota por $R_g^\delta(f)$ [13]. Como f actúa como límite de propagación, llega un momento donde la dilatación no produce ninguna variación en la señal, es decir, se trata de una apertura algebraica de la imagen máscara. En la siguiente figura se muestra un ejemplo de reconstrucción por dilatación en una señal de una dimensión:

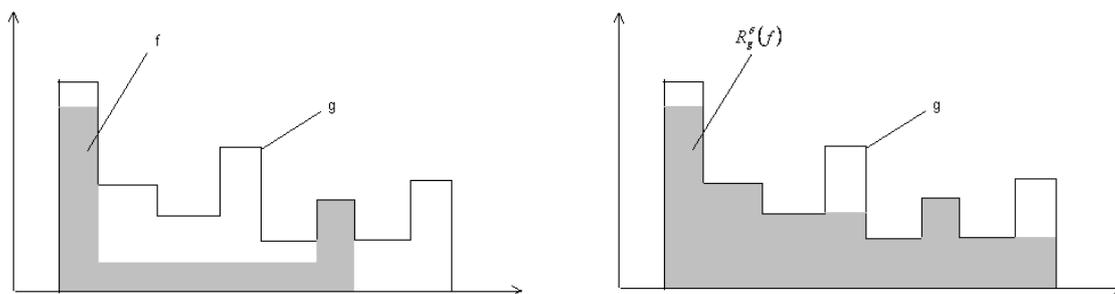


Figura 4.14. Ejemplo dilatación por reconstrucción de g sobre f

Particularizando para el algoritmo, se hace crecer la imagen $I_{fuerza-borde}$ usando como máscara la imagen I_{class} . En concreto, como la imagen que hace de máscara tiene que ser mayor o igual que la imagen marcador y en este caso en I_{class} no tiene por qué estar a uno todos los píxeles que salgan en la imagen de bordes, I_{fuerza_borde} , -porque como se verá en el siguiente capítulo, dependiendo de la imagen en cuestión, se puede detectar en esta última el borde de la retina, lo cual no se marca en la imagen resultado de la clasificación a uno-, se elige como máscara la imagen resultado de realizar una operación AND (lógica) entre ambas. Nuevamente, el resultado es una matriz binaria, donde se tiene a uno los píxeles que, hasta el momento, el algoritmo considera como pertenecientes a un exudado duro.

Hasta ahora, se han eliminado los falsos positivos correspondientes a otras lesiones que no tienen los bordes muy marcados, como sí ocurre en el caso de los exudados duros.

Llegado a este paso, sólo falta eliminar el disco óptico, que tiene atributos similares a los EDs en brillo, color y contraste. Se realiza una detección automática del disco óptico para enmascarar hacia fuera de la región y quitarla del resultado final, tal y como se explicó previamente en el Apartado 4.3.

Enmascarando la región del disco óptico al resultado obtenido en la dilatación, se obtiene el resultado final y definitivo del algoritmo, en el que ya sólo aparecen exudados duros, descartando otras lesiones amarillentas y el disco óptico. En la figura 4.15 se ha representado sobre la imagen original los exudados duros detectados marcados en verde.

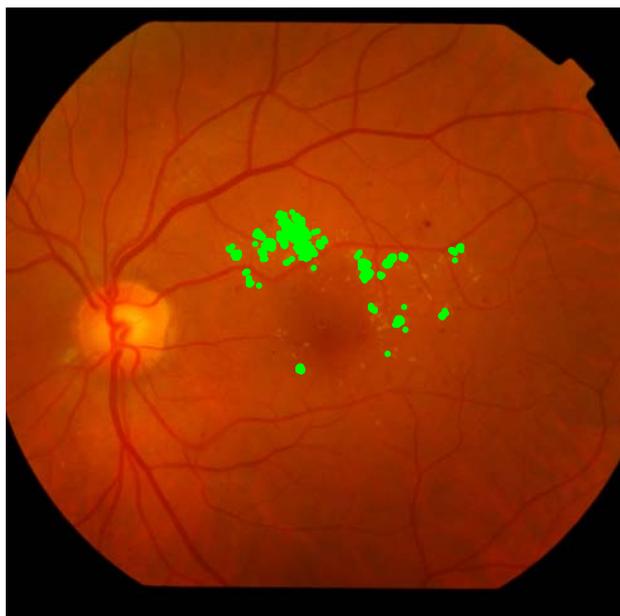


Figura 4.15. Resultado final del algoritmo

Hasta aquí la descripción de todas las etapas del algoritmo para detectar EDs. Como ayuda para seguir la explicación, se ha mostrado cada paso aplicado sobre una imagen en concreto. En el próximo capítulo, se mostrarán los resultados obtenidos sobre un conjunto de imágenes reales de retina de diferentes características para comprobar el funcionamiento del algoritmo.