

Oculis: Graduación automática de la hiperemia ocular

Pablo García Fernández
pablo.garcia.fernandez2@rai.usc.es

Objetivo—El objetivo de este informe es recoger el trabajo desarrollado por el alumno durante la ejecución de la práctica 1 de la asignatura *Instrumentation and Processing for Biomedical Applications*, así como los resultados obtenidos.

I. INTRODUCCIÓN

Uno de los síntomas que pueden aparecer en el ojo humano como consecuencia de una patología (p. ej. alergia, xeroftalmía, infecciones oculares, etc.) es la hiperemia en la conjuntiva bulbar [1]. Esto ocurre cuando la sangre queda atrapada en la zona afectada produciéndose, entre otros, signos visibles como un incremento de la rojez.

El objetivo de este trabajo es desarrollar, a través de técnicas básicas de procesamiento de imágenes, una metodología que permita cuantificar de forma automática el nivel de rojez de un ojo.

La metodología propuesta recibe como entrada una imagen RGB. Esta imagen es segmentada para extraer automáticamente la parte del ojo a ser analizada, es decir, la conjuntiva. Para ello un conjunto de *thresholdings* sobre diferentes espacios de color son aplicados sobre la imagen RGB [2]. Una suma ponderada de los resultados permite distinguir que píxeles forman parte de la máscara final y cuáles no. A continuación, las regiones de interés son depuradas vía filtros de media y aperturas morfológicas.

Extraída la conjuntiva, dos aproximaciones distintas (detector Canny [3] y *adaptive local thresholding*) son aplicadas sobre ella para localizar automáticamente los vasos sanguíneos.

Por último, una serie de características son extraídas de la imagen procesada. Entre ellas una estimación del área ocupada por los vasos, el nivel de rojez general de la conjuntiva y el grado de rojez de los vasos. Estas características son empleadas para la definición manual de reglas de graduación que permiten clasificar la hiperemia ocular.

Para el desarrollo del trabajo se dispone de un conjunto de imágenes de diferentes grados de hiperemia ocular. Para evaluar la precisión en la identificación de la conjuntiva, se han segmentado las 10 imágenes que forman parte de este conjunto de datos tal y como se recoge en la “Fig. 1”. Estas máscaras reales serán comparadas con las máscaras generadas por la metodología empleando la métrica IoU

Para evaluar la calidad de la clasificación se empleará el acierto.

II. SEGMENTACIÓN

El primer paso de la metodología desarrollada consiste en identificar la región del ojo a ser analizada. Partiendo de las

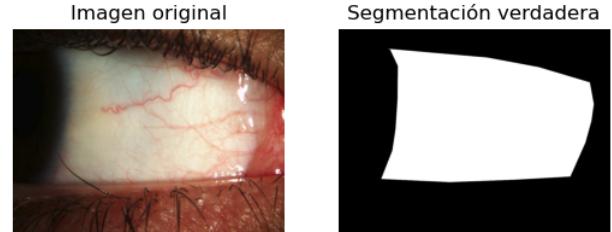


Fig. 1. Segmentación verdadera de una de las imágenes del *dataset*. Se utilizará para evaluar la calidad de las segmentaciones de la metodología.

ideas propuestas en [2] se ha optado por seguir un enfoque basado en *thresholdings* de la imagen de entrada en diferentes espacios de color. Este enfoque es rápido y ofrece buenos resultados. Incluso en pacientes no sanos, se espera que la conjuntiva sea más clara que el resto de los elementos del ojo (piel, pestañas, párpados, etc.). En esta línea se han probado 3 aproximaciones incrementales cada una de las cuales soluciona problemas de la aproximación anterior.

A. Aproximación 1

Debido a la gran variabilidad de las posibles imágenes de entrada, no existe un solo *thresholding* que funcione bien en todos los casos. Por este motivo, se han implementado 7 *thresholding* distintos (todos ellos extraídos de [2]). Concretamente:

- **Mtg.** *Thresholding* en el canal G de una imagen BGR usando la media de la intensidad de toda la imagen.
- **Mtg2.** *Thresholding* en el canal G de una imagen BGR usando la media del tronco central del canal G. El tronco central se construye dividiendo la imagen en tres partes horizontales. De esta forma se espera tener en cuenta una mayor cantidad de píxeles de la conjuntiva para la obtención del umbral.
- **Mtv.** *Thresholding* en el canal V de una imagen HSV usando la media del canal V.
- **Mtl.** *Thresholding* en el canal L de una imagen L*a*b* usando la media del canal L.
- **Mts.** *Thresholding* en el canal S de una imagen TSL usando la media del canal S
- **Mts2.** *Thresholding* en el canal S de una imagen TSL usando la media del canal S y considerando únicamente píxeles no rojos. Se corrige el nivel de rojo para paliar el efecto que tiene en la segmentación los casos de hiperemia más graves.

- **Mmo.** *Thresholding* en el canal G de una imagen BGR usando la media de la intensidad de toda la imagen. Seguido de aberturas morfológicas (10 iteraciones con *kernel* cuadrado 3x3).

Cada *thresholding* genera una máscara binaria de 1 y 0. Las 7 máscaras son combinadas mediante una simple suma en una nueva máscara *m*. Finalmente *m* es binarizado en función de un parámetro *n* (ajustado experimentalmente). Los píxeles mayores o iguales a *n* de *m* se ponen a 1; el resto a 0. Con esto lo que conseguimos es que un píxel solo forme parte de la máscara final si al menos *n* máscaras parciales así lo consideran.

Los resultados de esta aproximación se recogen en la “Tab. I”

TABLE I
RESULTADOS APROXIMACIÓN 1

mIoU	0.75
------	------

Aunque son aceptables, si los desglosamos para cada una de las imágenes observamos como esta aproximación falla para los casos más graves de hiperemia, donde los vasos sanguíneos son confundidos con el fondo “Fig. 2”.

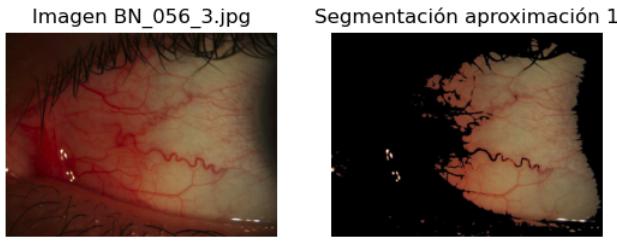


Fig. 2. La aproximación 1 falla para los casos más severos de hiperemia

Para solucionar estos problemas aparece la aproximación 2.

B. Aproximación 2

En los casos más severos de hiperemia los *thresholding* mtg, mtg2, mts y mmo tienden a confundir los vasos con el fondo; mientras que los *thresholding* mtv mtl y mts2 generan buenas máscaras “Fig. 3”.

Para solucionar este problema se optó por diferenciar dos situaciones. En función del nivel de rojo de la hiperemia (*rojo*), calculado como la media del tronco central del canal *a** de la imagen *L*a*b**, tenemos:

$$\begin{cases} m = 3mtv + 2mtl + 2mts2 & Si \ rojo > L_R \\ m = mtg + mtg2 + \dots + mmo & Si \ rojo \leq L_R \end{cases} \quad (1)$$

Donde *L_R* es un parámetro ajustado experimentalmente a 145.

De esta forma lo que se consigue es que en los casos más severos de hiperemia, la máscara final solamente tenga en cuenta aquellas máscaras parciales que logran un buen desempeño (mtv, mtl y mts).

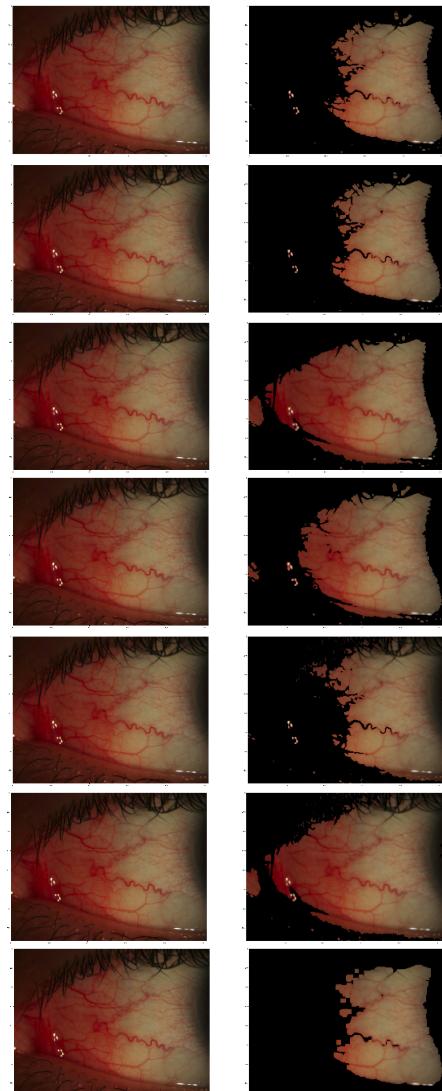


Fig. 3. De arriba a abajo mtg, mtg2, mtv, mtl, mts, mts2 y mmo. Se observa como mtg, mtg2, mts y mmo fallan en los casos más severos de hiperemia.

En “Fig. 4” se observa como con esta nueva aproximación, desaparecen los problemas al procesar imágenes muy rojas.

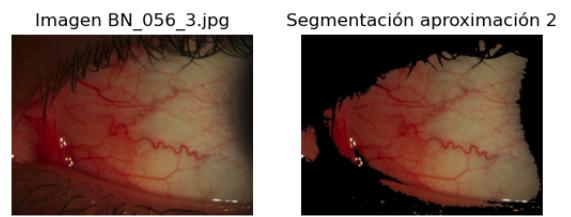


Fig. 4. A diferencia de la aproximación 1, ahora los casos severos de hiperemia son correctamente segmentados

Los resultados de procesar todas las imágenes con esta aproximación se recogen en la “Tab. II”

Aunque mejoran 0.02 puntos respecto a la aproximación anterior, si desglosamos los resultados para cada una de las

TABLE II
RESULTADOS APROXIMACIÓN 2

mIoU	0.77
------	------

imágenes que conforman el *dataset* observamos como esta aproximación aún presenta dos deficiencias:

- 1) Las imágenes en las que se incluyen partes de la piel de alta luminosidad generan “huecos” en la máscara final “Fig. 5”.
- 2) Las imágenes que incluyen vasos gruesos y rojos generan oclusiones en la conjuntiva “Fig. 5”.

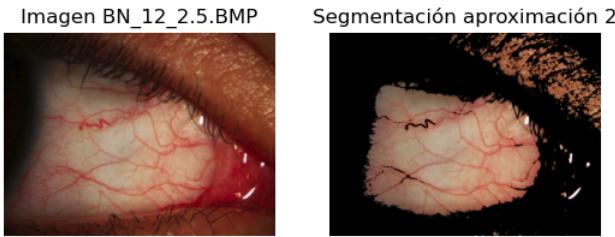


Fig. 5. En los bordes exteriores observamos los defectos en la segmentación por perturbaciones de piel clara (1). En la conjuntiva observamos los defectos por perturbaciones de vasos muy rojos (2).

C. Aproximación 3

Para solucionar estos problemas surge la aproximación 3. La idea subyacente es aplicar sobre la máscara final m un conjunto de filtros de media y operaciones morfológicas de apertura. Concretamente:

- N_MEDIAN filtros de media con *kernel* de tamaño 17×17 son aplicados sobre m para reducir la influencia de 0's aislados en la conjuntiva. Con esto se persigue eliminar el problema 2) difuminando los vasos detectados como fondo. N_MEDIAN es un parámetro ajustado experimentalmente a 4.
- N_A aperturas morfológicas con *kernel* elíptico de tamaño 23×23 son aplicados sobre m para eliminar los 1's fuera de la conjuntiva. Con esto se persigue eliminar el problema 1) “tapando” los huecos de m . N_A es un parámetro ajustado experimentalmente a 6.

Los resultados de esta aproximación se recogen en “Tab. III”

TABLE III
RESULTADOS APROXIMACIÓN 3

mIoU	0.83
------	------

Mejoran 0.06 puntos respecto a la aproximación anterior, alcanzando un nivel de precisión suficiente para lo que se quiere lograr con este trabajo. En “Fig. 6” se muestran las segmentaciones de las 10 imágenes del *dataset*.

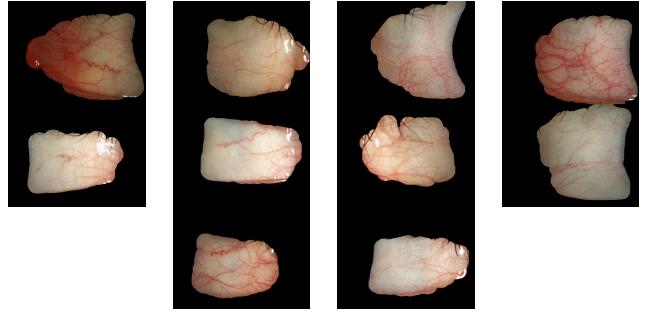


Fig. 6. Segmentación de las 10 imágenes que integran el *dataset*.

D. Eliminación de brillos

Para eliminar los brillos de la conjuntiva causados por reflejos de luz se optó por emplear un *thresholding* binario sobre el canal L^* de la imagen $L^*A^*B^*$. El valor del umbral seleccionado experimentalmente fue 240. El resultado de este proceso es dilatado vía operaciones morfológicas (5 iteraciones con *kernel* 3×3) para abarcar el área completa del brillo (el *thresholding* solo localiza los puntos centrales de máxima intensidad).

E. Enhancement techniques

Por ultimo, se valoró la aplicación de distintas técnicas de mejora de imágenes para obtener unos resultados más robustos. Lamentablemente, ninguna de las técnicas consideradas supuso una mejor sustancial “Tab. IV”.

TABLE IV
RESULTADOS enhancement techniques

Técnica	mIoU
White patch	0.83
Ecualizado	0.80
Gamma	0.79

III. LOCALIZACIÓN DE VASOS

Para la localización de los vasos sanguíneos de la conjuntiva se emplearon dos aproximaciones distintas. La primera de ellas consiste en encontrar los bordes de los vasos sanguíneos por medio del detector Canny. La segunda consiste en segmentar el vaso completo a través de un *thresholding* local y adaptativo.

A diferencia del apartado anterior, en este caso la evaluación de los resultados de la localización de los vasos se realizará de manera cualitativa (no se dispone del tiempo necesario para segmentar a mano las venas del *dataset*)

A. Aprox. 1: Canny

Se empieza ecualizando el resultado de la segmentación. Así, se consigue solventar el desafío que suponen las diferentes condiciones de iluminación y se mejoran los resultados.

A continuación la imagen es suavizada con un filtro gaussiano de tamaño 13×13 . Finalmente el detector Canny es aplicado para detectar las zonas de la imagen con valores

elevados en su primera derivada, es decir bordes (puntos de transición entre intensidades claras y oscuras). Los umbrales bajo y alto que establecen los márgenes de actuación de Canny fueron ajustados a 5 y 30 respectivamente.

En la “Fig. 7” se muestran los resultados obtenidos con esta aproximación.

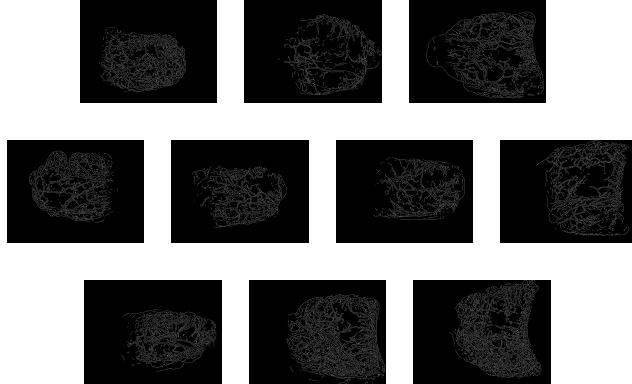


Fig. 7. Localización de todos los vasos a través de Canny.

B. Aprox. 2: Thresholding local adaptativo

Al igual que antes se empieza ecualizando el resultado de la segmentación. A continuación, una ventana de tamaño 21x21 se desplaza sobre la imagen segmentada. En cada movimiento la siguiente función es evaluada para determinar si el píxel forma parte o no del vaso:

$$\begin{cases} 1 & a_p \leq \text{media}(a) + k_1 \wedge a_p \geq \text{media}(a) - k_1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

donde a es el vecindario 21x21 del píxel central a_c . De esta forma únicamente son considerados vasos aquellas partes de la imagen que mantienen una intensidad más o menos constante en su vecindario (propiedad que cumplen las venas). k fue ajustado experimentalmente a 5.

Al producto de este *threshold* se le aplican dos suavizados de mediana con filtros de tamaños 5x5 y 3x3 para eliminar parte del ruido introducido en el proceso.

Los resultados de esta aproximación se recogen en la “Fig. 8”

C. Comparativa entre aproximaciones

Ambos métodos están integrados en el proyecto desarrollado. No se puede afirmar con rotundidad que uno es mejor que otro, pues cada uno aporta cosas diferentes. Concretamente:

- **Aprox. 1:** Se obtienen únicamente los bordes de los vasos. Es un método muy rápido (x3.25 respecto aprox. 2). La precisión sobre venas finas y desarticuladas es mayor.
- **Aprox. 2:** Se obtienen todos los píxeles que integran los vasos, esto permite estimar el área y demás características con mayor precisión. Es más lento. Menor precisión sobre vasos pequeños.

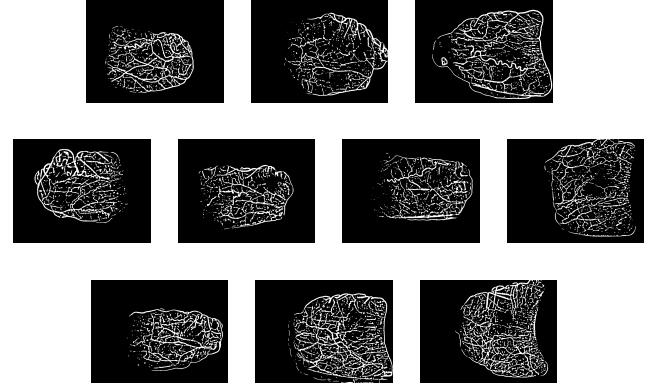


Fig. 8. Localización de todos los vasos a través de *thresholding* local.

IV. EXTRACCIÓN DE CARACTERÍSTICAS

La metodología propuesta finaliza extrayendo un conjunto de características de la imagen procesada; y utilizando las mismas para la clasificación del grado de severidad de la hiperemia.

Concretamente 2 características son computadas: el área ocupada por los vasos y el nivel de rojez de las venas. Su definición es la que sigue. En ambas ecuaciones n, m son las dimensiones de la región de interés (resultado de la segmentación) y $Venas$, la máscara binaria resultante de la localización de los vasos.

$$Area = \frac{\sum_i^n \sum_j^m V_{i,j}}{n \cdot m} \quad \forall V_{i,j} \in Venas \quad (3)$$

donde $V_{i,j}$ son los píxeles que forman parte de $Venas$ y se corresponden con un vaso.

$$Rojez = \frac{\sum_i^n \sum_j^m A_{i,j} \cdot Venas_{i,j}}{\sum_i^n \sum_j^m Venas_{i,j}} \quad (4)$$

donde $A_{i,j}$ son los valores de intensidad del canal A* de el espacio de color L*A*B*

V. CLASIFICACIÓN DEL GRADO DE SEVERIDAD DE LA HIPEREMIA

Inicialmente se probó a realizar la clasificación considerando únicamente el área de ocupación de los vasos. Sin embargo, esta aproximación fallaba para casos de hiperemia donde las venas ocupaban un área menor con un rojo muy intenso (p. ej. BN_056_3.jpg). En estos casos los oftalmólogos tienden a graduar la hiperemia en niveles medios/altos, pero el sistema predecía niveles leves (por la baja área). Para incluir esta apreciación se incluyó la característica asociada a la rojez de las venas. De esta forma se cubren las dos principales manifestaciones de la hiperemia: ocupación y rojez.

En “Fig. 9” se observa el modelo construido a mano como un árbol de clasificación:

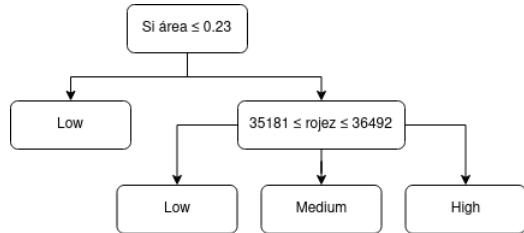


Fig. 9. Modelo de clasificación para localización con local *thresholding*

VI. DISCUSIÓN DE RESULTADOS

En “Fig. 10” se observa el resultado de aplicar la metodología propuesta sobre las 10 imágenes del dataset. Destacar que la aproximación seguida para la localización de los vasos es la del *thresholding* local. El acierto obtenido es del 100%

REFERENCES

- [1] Red eye. (17 de Noviembre de 2021). En wikipedia. [https://en.wikipedia.org/wiki/Red_eye_\(medicine\)](https://en.wikipedia.org/wiki/Red_eye_(medicine))
- [2] María Luisa Sánchez Brea. (2017). *Automatic grading of ocular hyperaemia using image processing techniques*.
- [3] John Canny. (1986). *A computational approach to edge detection*. En Pattern Analysis and Machine Intelligence, IEEE Transactions on, no. 6, pp. 679–698, 1986.

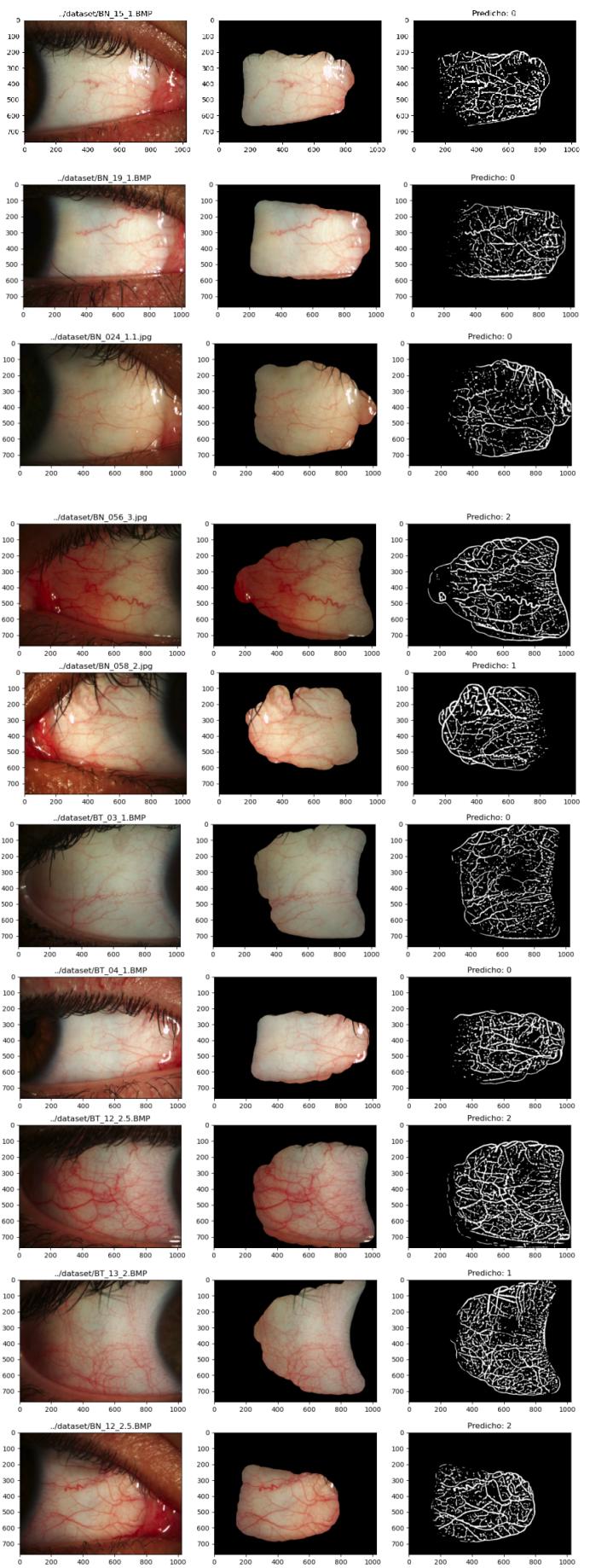


Fig. 10. Resultados de la metodología.