

# Clasificador de pigmentación de la ballena azul aplicando el método SIFT en medios no controlados

Rosa I. Ramos-Arredondo<sup>1</sup>, Blanca E. Carvajal-Gómez<sup>2</sup>,  
Francisco J. Gallegos-Funes<sup>3</sup>, Diane Gendron-Laniel<sup>4</sup>

<sup>1</sup> Escuela Superior de Cómputo, Instituto Politécnico Nacional, SEPI-ESCOM,  
México D.F.

<sup>2</sup> Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas,  
Instituto Politécnico Nacional, México D.F.

<sup>3</sup> Escuela Superior de Ingeniería Mecánica y Eléctrica, Instituto Politécnico Nacional,  
ESIME SEPI-Electrónica, México D.F.

<sup>4</sup> Centro Interdisciplinario de Ciencias Marinas, Instituto Politécnico Nacional,  
La Paz, B.C.S., México

alesija@gmail.com, {becarvajal, fgallegosf, dgendron}@ipn.mx

**Resumen.** Clasificar las imágenes recolectadas por los especialistas marinos en medios no controlados de la tonalidad de pigmentación de la piel de la ballena azul, haciendo uso del método de foto-identificación; es necesario aplicar un método para la extracción de características que identifique a cada imagen con el método SIFT (Scale-invariant feature transform). Se analiza el rendimiento del clasificador y el desarrollo del método en conjunto; valorando la precisión, la sensibilidad, la especificidad y su rango de falso positivos. Al clasificar 500 muestras de cinco tonalidades de pigmentación (clara, clara-moteada, franja, balanceada y oscura), se obtuvo un 81.80 % de precisión para la tonalidad balanceada y 80.40 % con la tonalidad clara-moteada, tomando 100 muestras de entrenamiento. Al aumentar el número de muestras de entrenamiento a 150, la tonalidad balanceada se mantiene con un 81.20% de precisión, la tonalidad clara-moteada aumentando su precisión a 81%. Al disminuir el número de muestras de entrenamiento a 50, la precisión de la tonalidad oscura fue del 79.80 % y 78.60% de la tonalidad balanceada.

**Palabras claves:** clasificadores, procesamiento de imágenes, extracción de características, ballena azul.

## 1. Introducción

La identificación de la ballena azul por medio del método de foto-identificación se basa en clasificar a la especie por medio de características representativas que las hacen únicas e irrepetibles entre otras especies. [1] Para el caso de la ballena azul, la

identificación de la tonalidad de la piel de la ballena azul como la forma de la aleta dorsal son características esenciales para la clasificación de las fotografías recolectadas en campo por los especialistas marinos. En el Laboratorio de Ecología de Cetáceos y Quelonios del Centro Interdisciplinario de Ciencias Marinas (CICIMAR) cuenta con un acervo de imágenes de la ballena azul, clasificadas por tonalidades de pigmentación y por la forma de la aleta dorsal; que han sido recolectados por más de 25 años. La recolección de las imágenes de la ballena azul en campo es complicado controlar las condiciones y el medio de exposición del individuo de estudio y las condiciones de su hábitat. Las condiciones generales son; estar a más de 100 m de distancias del individuo de estudio y recolectar imágenes en condiciones climáticas favorables de iluminación.

El proceso de clasificar las imágenes por la tonalidad de pigmentación y la forma de la aleta dorsal por parte de los especialistas marinos es una tarea, en donde se involucran de al menos tres especialistas con experiencia en la especie del individuo de estudio para determinar de forma veraz y sin dar a equivocaciones. Es necesario el desarrollo de métodos y herramientas en la clasificación de las imágenes, además que se recolecta un gran volumen de imágenes en cada avistamiento y no todas son candidatas para ser utilizadas en la identificación.

En el estudio de las imágenes a clasificar se comprobó que el factor de ser en medios no controlados la exposición de las imágenes; provocara que el elemento a identificar se encontrará en diferentes proporciones, posición y escalas. El empleo del método SIFT para la extracción de características de las imágenes es por ser robusto a transformaciones locales y globales, como también a perturbaciones de cambios en escalamiento, rotación y posicionamiento. Elementos a considerar en el momento de clasificar una gran cantidad de imágenes.

El presente trabajo de investigación se encuentra dividido en secciones; en la primera sección métodos y materiales; se enfocará en exponer los métodos que se necesitarán para la clasificación de los tonos de pigmentación de la ballena azul y la aplicación del método SIFT, como los materiales involucrados para la realización de los experimentos. En la segunda sección resultados, se muestran los datos obtenidos a partir de los experimentos, mostrados en las tablas. En la tercera sección, se discuten los resultados obtenidos y se enfoca en efectuar las comparativas entre los procedimientos experimentados.



**Fig. 1.** Diagrama de bloques del proceso general para clasificar a la ballena azul por su tono de pigmentación.

## 2. Métodos y materiales

En esta sección nos enfocaremos en el desarrollo de los métodos y materiales para el clasificador de las imágenes por tonos de pigmentación de la piel de la ballena azul. En la Figura 1, se muestra el diagrama de bloques de las etapas y el flujo que nos permitirá clasificar a las imágenes de la ballena azul por sus tonos de pigmentación.



**Fig. 2.** Imágenes del acervo del CICIMAR de la ballena azul, a) Imagen mostrando el flanco izquierdo de la ballena azul, b) imagen de la ballena azul en donde se muestra el lomo sin su aleta dorsal, c) Imagen mostrando una porción reducida del lomo de la ballena azul

### 2.1 Imágenes a clasificar

En el CICIMAR cuenta con un acervo de imágenes previamente clasificadas por la forma de la aleta dorsal y sus tonos de pigmentación de la piel de la ballena azul. En este trabajo de investigación nos centramos en la clasificación de las imágenes de la ballena azul por su tono de pigmentación y para el caso de estudio tomaremos cinco clases (clara, clara-moteada, franja, balanceada y oscura) [2]. Se tomarán un total de 846 imágenes y su distribución entre clases es; clara (101), clara moteada (304), franja (116), balanceada (174), oscura (151). Información adicional, la clase clara-moteada y

la clase franja presenta un patrón además de un tono de pigmentación. Esta consideración no altera el estudio de la investigación, ni se añadirá alguna etapa extra.

## 2.2 Procesamiento de las muestras

Las imágenes del acervo a utilizar no tiene una dimensión concreta y al ser recolectadas la porción de la ballena azul en la imagen es variado como su posición, su ángulo y en algunos casos la distancia en la que fue tomada.

**Selección de la región de interés.** La selección de la región de interés (ROI, region of interest) de las imágenes en la ballena azul es para reducir el número píxeles al analizar, como el número de características a clasificar del tono de pigmentación en conjunto. Por tal motivo se reduce las tareas computacionales al analizar la imagen completa. Se tomará una región cuadrada de 100 píxeles cuadrados en el lomo de la ballena azul y no próxima de la aleta dorsal para evitar un análisis adicional como; segmentación, realce de píxeles entre otros.

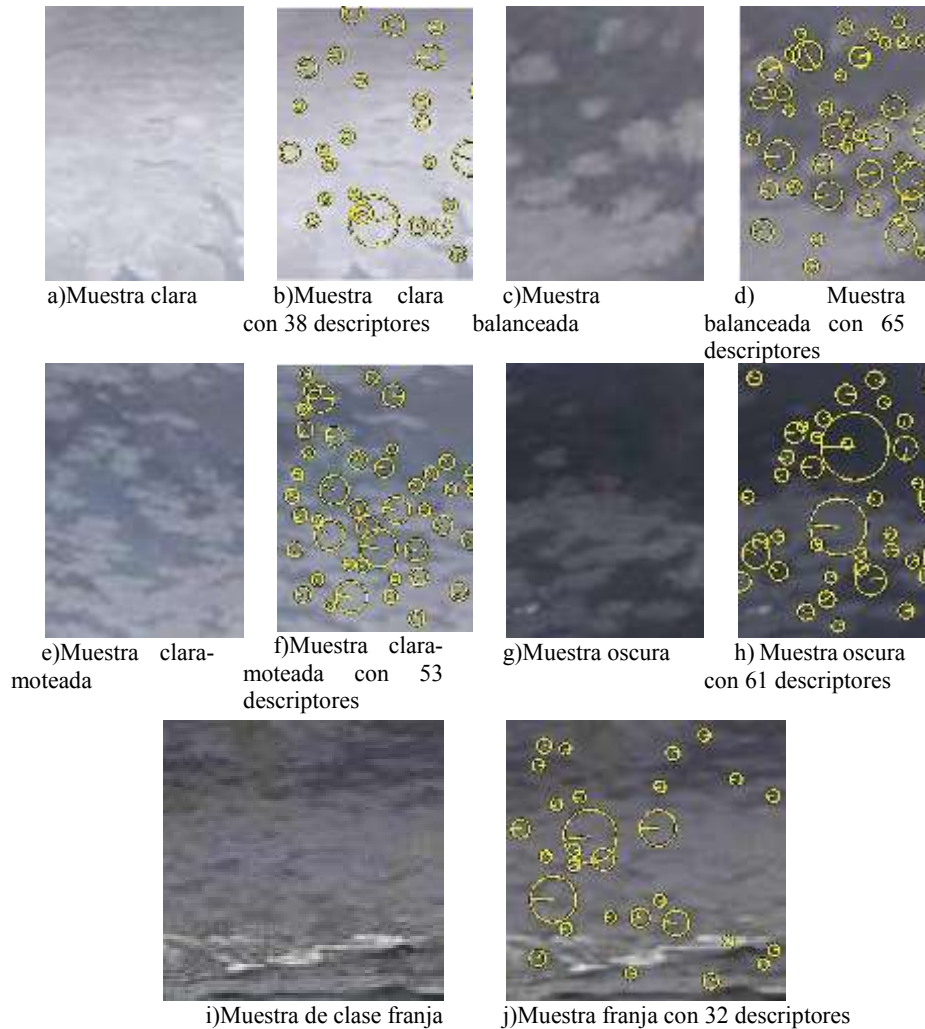


**Fig. 3.** Selección de la región de interés a analizar de la ballena azul

Las regiones de interés de las imágenes del acervo se ocuparán para su clasificación y una porción de estas, se utilizará para la etapa de entrenamiento. Una muestra puede describirse como  $I$  en donde,  $f(x,y)$  representa el nivel de intensidad en las coordenadas  $(x,y)$ . Véase en la ecuación 1. La región de interés estará acotada en una región cuadra de 100 píxeles cuadrados.

$$I = f(x,y). \quad (1)$$

**Extracción de características con SIFT.** Las muestras a clasificar como las muestras para un previo entrenamiento, se les extraerá características, que en la etapa de clasificación se ocupara para identificar. Es cuando el método SIFT es utilizado y nos proporcione un arreglo de descriptores (descriptor) y puntos claves (keypoints). El algoritmo consiste en cuatro etapas: 1) selección de escala-espacio de picos (scale-space peak selection), b) localización de punto clave (keypoint localization), c) asignación de orientación (orientation assignment), d) descriptores de punto clave (keypoint descriptor).[3] El aplicar este método a las muestras a clasificar es porque nos ubica puntos de interés, además de ser un algoritmo invariante a cambio de escala, rotación de la imagen y pequeños cambios de iluminación; condiciones que se da en la recolección de imágenes de la ballena azul en medio no controlados.



**Fig. 4.** Muestras de cada uno de los tonos de pigmentación a estudiar y su representación de los descriptores de cada uno, como su magnitud y ángulo. a),c),e), g),i) muestras de los tonos de pigmentación y b),d),f),h),j) su representación de los descriptores.

### 2.3 Clasificación

Todas las muestras (imágenes que clasificar e imágenes que se ocuparán en el entrenamiento) después de obtener la extracción de características se compone por una estructura (frame) y descriptores. El número de descriptores es variable para cada imagen. La correspondencia entre los descriptores de las imágenes en la clasificación es obteniendo la distancia euclidiana entre los descriptores de la imagen más cercana y la siguiente, obteniendo solo el número de puntos de encuentro. La clasificación de las

imágenes de la ballena azul en la literatura se encuentra el estudio por medio de la forma de la aleta dorsal, en donde se emplea un clasificador supervisado, con una previa selección de muestras para el entrenamiento [4].

**Entrenamiento.** Las muestras que se seleccionan para el entrenamiento son imágenes representativas de cada tono de pigmentación o clase (termino que se ocupará en el desarrollo del clasificador). La selección de muestras de entrenamiento se ocupa como punto de referencia para la clasificación de muestras a identificar. Las muestras a clasificar y las de entrenamiento después del proceso de extracción de características empleando el método SIFT, tiene un vector de características y tomaremos el vector descriptor  $\#d_i = |d_i|$ , donde  $i$ , es el número de elementos que lo componen. Donde  $\bar{C}_j$ , es el valor representativo por cada  $j$  clases; las clases utilizadas son clara, clara moteada y oscura. Véase en la ecuación 2.

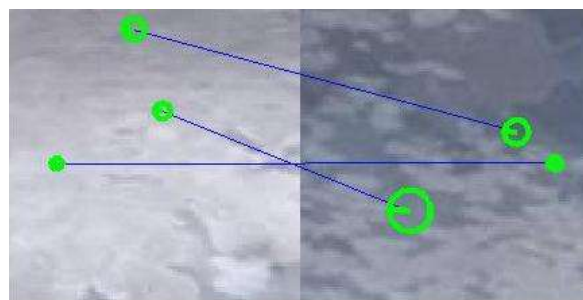
$$\bar{C}_j = \frac{\sum_{i=1}^n |d_i|}{n}. \quad (2)$$

El vector de entrenamiento  $E$  que se empleara en la clasificación de muestras estará compuesto por  $k=5$  valores, véase en la ecuación 3.

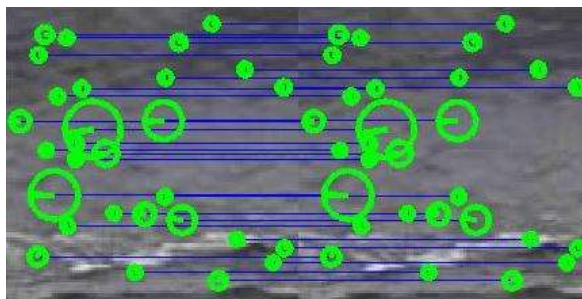
$$E = (\bar{C}_1, \bar{C}_2, \dots, \bar{C}_k). \quad (3)$$

Se tomara el mismo número de muestras de entrenamiento de cada clase para evitar una sobre clasificación. Además, de tomar un número pequeño con relación al número de muestras a clasificar. De las muestras de entrenamiento solo se estudiara sus descriptores.

**Conteo de encuentros.** Las muestras a clasificar se analizaran por medio de su vector descriptor  $\#d_i$ , el cual se comparará su distancia con los vectores descriptores de las muestras de entrenamiento de cada una de las clases. Posteriormente se clasificará tomando el valor máximo con relación al vector de entrenamiento  $E$ ; proporcionando la pertenencia a una de las clases (tonos de pigmentación de la piel de la ballena azul). En la figura 5 se ilustra los puntos de encuentro entre muestras de diferentes tonos de pigmentación y de iguales tonos.



a) Puntos de encuentro entre una muestra clara y clara moteada



b) Puntos de encuentros entre dos muestras franja

Fig. 5. Ejemplo de puntos de encuentro entre dos muestras.

### 3. Resultados

Se ocuparon cien muestras de cada tono de pigmentación de la piel de la ballena azul para el estudio (clara, moteada, franja, balanceada y oscura), analizando un total de 500 muestras a clasificar. Para las muestras de entrenamiento se analizaron con 50,100 y 150 muestras, tomando de forma equitativa entre los tonos de pigmentación de las muestras a clasificar.

En la Tabla 1, se muestra la clasificación de 500 muestras con relación de 50 muestras de entrenamiento (10 muestras de cada clase), dando la clase clara el mayor número de clasificación en comparación de las clases moteada y balanceada.

Tabla 1. Clasificación de cinco clases utilizando cincuenta muestras de entrenamiento.

Tono de pigmentación (Número de muestras)	Clasificación de Clases				
	Clara	Moteada	Franja	Balanceada	Oscura
Clara(100)	59	8	21	2	10
Moteada(100)	45	13	27	7	8
Franja(100)	43	7	39	4	7
Balanceada(100)	51	8	17	13	11
Oscura(100)	40	3	15	7	35

En la Tabla 2, se muestran las matrices de confusión de cada una de las clases, proporcionando la relación entre las 100 muestras a clasificar de forma supervisada y las 400 muestras de otras clases a analizar.

Tabla 2. Matriz de confusión del clasificador de cinco clases utilizando cincuenta muestras de entrenamiento

	Clara		Moteada		Franja		Balanceada		Oscura	
	Condición		Condición		Condición		Condición		Condición	
	+	-	+	-	+	-	+	-	+	-
Prueba Positiva	59	179	13	26	39	80	13	20	35	36
Prueba Negativa	41	221	87	374	61	320	87	380	65	364
Total	100	400	100	400	100	400	100	400	100	400

En la Tabla 3, se muestra la clasificación de 500 muestras con relación de 100 muestras de entrenamiento (20 muestras de cada clase), proporcionando la clase balanceada caso nulos en la clasificación con otras clases.

**Tabla 3.** Clasificación de cinco clases utilizando cien muestras de entrenamiento.

Tono de pigmentación (Número de muestras)	Clasificación de Clases				
	Clara	Moteada	Franja	Balanceada	Oscura
Clara(100)	62	2	18	0	18
Moteada(100)	36	10	32	1	21
Franja(100)	31	2	51	0	16
Balanceada(100)	37	4	21	10	28
Oscura(100)	30	0	20	0	50

En la Tabla 4 se muestran las matrices de confusión de cada una de las clases. Corroborando el caso de la clase balanceada y moteada como las clases con menor número de errores en la clasificación supervisada.

**Tabla 4.** Matriz de confusión del clasificador de cinco clases utilizando cien muestras de entrenamiento.

	Clara		Moteada		Franja		Balanceada		Oscura	
	Condición		Condición		Condición		Condición		Condición	
	+	-	+	-	+	-	+	-	+	-
<b>Prueba Positiva</b>	62	134	10	8	51	91	10	1	50	83
<b>Prueba Negativa</b>	38	266	90	392	49	309	90	399	50	317
<b>Total</b>	100	400	100	400	100	400	100	400	100	400

En la Tabla 5, se muestra la clasificación de 500 muestras con relación de 150 muestras de entrenamiento (30 muestras de cada clase), proporcionando la clase balanceada caso nulos en la clasificación con otras clases.

**Tabla 5.** Clasificación de cinco clases utilizando ciento cincuenta muestras de entrenamiento.

Tono de pigmentación (Número de muestras)	Clasificación de Clases				
	Clara	Moteada	Franja	Balanceada	Oscura
Clara(100)	69	1	17	0	13
Moteada(100)	44	13	29	0	14
Franja(100)	44	1	41	0	14
Balanceada(100)	48	4	17	7	24
Oscura(100)	32	2	16	1	49

En la Tabla 6, se muestran las matrices de confusión de cada una de las clases. Corroborando el caso de la clase clara con el mayor número de errores en la clasificación supervisada, en cambio la clase balanceada y moteada siguen presentando el menor número de errores.



**Tabla 6.** Matriz de confusión del clasificador de cinco clases utilizando ciento cincuenta muestras de entrenamiento.

	Clara		Moteada		Franja		Balanceada		Oscura	
	Condición		Condición		Condición		Condición		Condición	
	+	-	+	-	+	-	+	-	+	-
<b>Prueba Positiva</b>	69	168	13	8	41	79	7	1	49	65
<b>Prueba Negativa</b>	31	232	87	392	59	321	93	399	51	335
<b>Total</b>	100	400	100	400	100	400	100	400	100	400

#### 4. Discusión

El estudio del rendimiento de la clasificación de la tonalidad de la pigmentación de la piel de la ballena azul hace uso de cuatro parámetros; verdadero positivo (VP) se prueba de forma positiva y la condición es positiva, falso positivo (FP) se prueba de forma positiva pero la condición es negativa, verdadero negativo (VN) se prueba de forma negativa dando negativa la condición y falso negativo (FN) se prueba de forma negativa dando positiva la condición. Estos parámetros se encuentran concentrados en la matriz de confusión de cada clase del clasificador. Son utilizados en las siguientes cuatro evaluaciones; especificidad (SP), sensibilidad (SN), la razón de falso positivos (FPR) y la precisión (ACC) [5-7].

En la Tabla 7, se concentra las evaluaciones de cada una de las clases tomando 50 muestras de entrenamiento. La clase con menor precisión es la clase clara como también en cuanto a sensibilidad. En cambio la clase moteada, balanceada y oscura superan el 77% de precisión.

**Tabla 7.** Rendimiento de la clasificación utilizando cincuenta muestras de entrenamiento.

	Clara	Moteada	Franja	Balanceada	Oscura
<b>Especificidad (%)</b>	84.35	81.12	83.98	81.37	84.84
<b>Sensibilidad (%)</b>	24.78	33.33	32.77	39.39	49.29
<b>FRP (%)</b>	15.64	18.87	16.01	18.62	15.15
<b>Precisión (%)</b>	56	77.40	71.80	78.60	79.80

En la Tabla 8, se concentra las evaluaciones de cada una de las clases tomando 100 muestras de entrenamiento. La clase moteada y balanceada obtiene una precisión arriba del 80%, la clase franja y oscura su precisión es del 72% y 73.40%.

**Tabla 8.** Rendimiento de la clasificación utilizando cien muestras de entrenamiento.

	Clara	Moteada	Franja	Balanceada	Oscura
<b>Especificidad (%)</b>	87.50	81.32	86.31	81.59	86.37
<b>Sensibilidad (%)</b>	31.63	55.55	35.91	90.90	37.59
<b>FRP (%)</b>	12.50	18.67	13.68	18.40	13.62
<b>Precisión (%)</b>	65.60	80.40	72	81.80	73.40

En la Tabla 9, se concentra las evaluaciones de cada una de las clases tomando 150 muestras de entrenamiento. La clase moteada y balanceada obtiene una precisión arriba del 80%, la clase franja y oscura su precisión es del 72% y 76.80%.

**Tabla 9.** Rendimiento de la clasificación utilizando ciento cincuenta muestras de entrenamiento.

	<b>Clara</b>	<b>Moteada</b>	<b>Franja</b>	<b>Balanceada</b>	<b>Oscura</b>
<b>Especificidad (%)</b>	88.21	81.83	84.47	81.09	86.78
<b>Sensibilidad (%)</b>	29.11	61.90	34.16	87.50	42.98
<b>FRP (%)</b>	11.78	18.16	15.52	18.90	13.21
<b>Precisión (%)</b>	60.20	81	72.40	81.20	76.80

## 5. Conclusión

En este trabajo de investigación, se llegó al análisis de cinco clases de tonos de pigmentación de la piel de la ballena azul (clara, clara-moteada, franja, balanceada y oscura), dando la clase balanceada la más alta en precisión independientemente del número de muestras en el entrenamiento del clasificador empleando el método SIFT en la extracción de características. Se observó que entre las clases moteada, franja y oscura se mantenía la precisión y se mejoró al aumentar el número de muestras de entrenamiento, con 100 y 150 muestras.

En el caso específico de la clase clara, el análisis nunca alcanza un porcentaje en precisión aceptable y fue la clase en donde se obtuvo un mayor número de errores de clasificación. La clase clara presenta características muy similares entre otras clase, por tal motivo se propone en futuros estudios; el reducir el número de clases a clasificar.

**Agradecimientos.** Los autores agradecen al Instituto Politécnico Nacional de México (IPN) y al Consejo Nacional de Ciencia y Tecnología de México (CONACYT) por la ayuda y el apoyo con número de proyecto 221284. Como la colaboración y apoyo otorgado por parte de la Dra. María Chantal Diane Gendron Laniel y a su equipo de profesionales colaboradores.

## Referencias

1. Randall S. Wells.: Identification Methods. In: Perrin, William F., Bernd Wursig, and J. G. M. Thewissen. (eds.) Encyclopedia of marine mammals, pp.593-596. Academic Press, Diego (2009)
2. Gendron, D., Ugalde De, A.: A new classification method to simplify blue whale photo-identification technique. Journal of Cetacean Research and Management, pp. 79-84. (2012)
3. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. In: International journal of computer vision, vol. 60.2, pp. 91-110. Kluwer Academic Publishers (2004)
4. Ramos-Arredondo, R. I., Carvajal-Gómez, B. E., Gallegos-Funes, F. J., Gendron-Laniel, D.: Multi-Spatial Classifier for Blue Whale Images using Photo-Identification Method. In: Castro Espinoza, F. (eds.) MICAI 2014: Advances in

- Artificial Intelligence, vol. 82, pp. 31-40. Center for Computing Research of IPN (2014)
5. Fathi, A., Naghsh-Nilchi, A.R.: Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation. *Biomedical Signal Processing and Control*. vol. 8, pp.71–80. (2013)
  6. Trejo-Salazar, D., Carvajal-Gómez, B. E., Gallegos-Funes, F. J.: Algoritmo de segmentación para imágenes reales para la clasificación de cetáceos. En: M. González-Mendoza y F. Castro-Espinoza (eds.). En: Congreso Mexicano de Inteligencia Artificial (COMIA 2013): Research in Computing Science. vol. 62, pp. 209-218, (2013).
  7. Fraz, M.M., Barman, S.A., Remagnino, P., Hoppe, A., Basit, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G.: An approach to localize the retinal blood vessels using bit planes and centerline detection. *Computer Methods and Programs in Biomedicine*, vol. 108, pp: 600–616, 2012