

VISIÓN POR COMPUTADORA PARA LA CLASIFICACIÓN Y DETECCIÓN DE FAUNA DE LA SIERRA NOR-ORIENTAL DE PUEBLA

CLASSIFICATION AND DETECTION OF FAUNA OF THE NOR- ORIENTAL MOUNTAIN RANGE OF PUEBLA

Luis Alberto Espejo Ponce

TecNM / Instituto Tecnológico Superior de Zacapoaxtla, México
jcensan@hotmail.com

Yareny Rivera García

TecNM / Instituto Tecnológico Superior de Zacapoaxtla, México
yarenyrivera@hotmail.com

José Miguel Méndez Alonso

TecNM / Instituto Tecnológico Superior de Zacapoaxtla, México
xtrem_jam@hotmail.com

Recepción: 31/octubre/2019

Aceptación: 9/diciembre/2019

Resumen

En el campo de la conservación de la fauna existen diversas técnicas para su estudio; como la estimación, captura, inventarios y monitoreo, mismas que, en algunos, casos dificultan y retrasan la cuantificación de una determinada especie. La presente investigación está dirigida al desarrollo de un sistema de clasificación de imágenes basado en algoritmos de visión por computadora para el monitoreo de fauna de la Sierra Nor-Oriental de Puebla. El sistema se basó en tres etapas para la identificación y clasificación; en primera instancia se realizó la extracción de características a un banco de imágenes representadas por tres especies, para establecer los puntos de interés donde se encuentra la información de las formas y patrones correspondientes, mediante el algoritmo SIFT, para después crear las fronteras de clasificación con el modelo Bag to Visual Words y obtener un diccionario de palabras visuales que se utilizó para generar un entrenamiento, como última etapa se realizaron pruebas con un conjunto de muestras distintas a las utilizadas en la generación del diccionario pero pertenecientes a las tres especies ,

como resultado el clasificador marco una tendencia hacia una de las especies generando falsos positivos en la clasificación general.

Palabras Clave: Algoritmos de Clasificación, Clasificación de Fauna, Visión por Computadora.

Abstract

In the field of wildlife conservation there are various techniques for its study; like the estimation, capture, inventories and monitoring, which, in some cases, hinder and delay the quantification of a certain species. This research is at the development of an image classification system based on computer vision algorithms for the monitoring of endemic fauna of the North-Eastern of Puebla. The system is based on three steps for identification and classification; First, the extraction of characteristics of the species to be identified is carried out, based on an image, to establish the points of interest where the information of the corresponding shapes and patterns is found, afterwards, the image characteristics is represented in vectors, finally to give way to the classification by the algorithms of K-Means or classification SVM, in which the search by similarity is carried out, ie, it is searched in a database of images that are visually similar to the image that is being classified, finally the image is tagged with the specie identified.

Keywords: *Active stretching, elongation, muscle proteins, passive stretching.*

1. Introducción

La conservación de la fauna en México establece un compromiso de Estado, dada la riqueza biológica que existe en el país, ejemplo de ello es que en México se alberga alrededor del 10% de las especies silvestres registradas en el mundo, ocupa el cuarto lugar en anfibios, segundo en mamíferos y primero en reptiles, posicionándolo como uno de los 17 países mega diversos [SEMARNAT, 2015].

Para garantizar la estabilidad en los diferentes ecosistemas es importante realizar evaluaciones que permitan conocer los cambios y así mantener una conservación optima de las especies. En México la Ley General de Vida Silvestre contempla las

leyes para controlar, conservar y manejar las diversas poblaciones de los ecosistemas [Cámara de Diputados del H. Congreso de la Unión, 2018].

Para lograr lo anterior es indispensable realizar estrategias apoyadas de herramientas y metodologías para conocer el estado de la población de la fauna, dentro de estas técnicas se encuentra el monitoreo, que es una herramienta que se utiliza para determinar los cambios en los ecosistemas, mediante la extracción de información que se obtiene a partir de capturar y marcar, conteo de huellas, madrigueras o identificación de excremento, así como el uso de foto trampas; esta última permite detectar especies sigilosas, estimar abundancia y densidad poblacional, así como conocer hábitos de comportamiento, su funcionamiento se basa en cámaras con sensores de movimiento y/o calor para disparar y tomar la fotografía, este tipo de foto trampas solo captura y en algunos casos envía la imagen mediante alguna tecnología de comunicación lo que hace necesaria la concentración de las imágenes para su interpretación [CONAFOR, 2017] [Gallina y López, 2019]. Existen diversas investigaciones enfocadas a la cuantificación y a la inspección de fauna; como el trabajo de Hernández *et al.* [2018], en el que establece un sistema para estimar la población de murciélagos en varias imágenes, esto lo realizan mediante la segmentación de imágenes donde se cree, puedan existir murciélagos, esto mediante el método de Otsu, así como la aplicación de filtros de mediana y filtros Gaussiano, en el trabajo se establece una comparación del conteo manual y mediante la técnica de visión por computadora, teniendo como resultado un conteo equiparable. En el trabajo realizado por Cohen *et al.*, 2015, establece un sistema integral para reconocer e identificar animales mediante la obtención de videos e imágenes en entornos controlados para realizar un posterior análisis, en su investigación establecen una arquitectura que contiene un módulo para la captura de imágenes y video, un clasificador Haar Cascade, un clasificador Local Binary Pattern, estos algoritmos son utilizados para detectar, identificar y realizar seguimiento de los animales, así como una red neuronal para validar los resultados de los clasificadores. El sistema fue probado en un entorno controlado en el que se estableció a la ardilla como animal objetivo, teniendo resultados favorables al detectar su presencia.

El presente trabajo se centra en la integración de algoritmos de visión por computadora y aprendizaje máquina para realizar una interpretación de las imágenes y permitir la clasificación de una fauna determinada, para lograr esto es necesario contar con un conjunto de imágenes de la especie de fauna que se desea detectar y clasificar, primero se realiza la extracción de características a cada una de las imágenes para conocer los puntos de interés y así identificar a la especie, para después crear mediante el modelo de *Bag of word* la representación de cada imagen como un histograma de clases y utilizar el algoritmo *Support Vector Machines* (SVM) para la clasificación.

Transformación de característica invariante de escala (SIFT)

SIFT es un algoritmo que permite la extracción de puntos clave en una imagen y después permite describirlos mediante un histograma orientado de gradientes, presentando invarianza a la escala, rotación e iluminación. Para obtener las descripciones SIFT fue propuesto para trabajar sobre imágenes en escala de grises, por lo que el vector de características de 128 elementos que define a cada pixel, contiene información sobre cómo se distribuyen los niveles de intensidad sobre los puntos de interés obtenidos [Lowe, 2004]. Por lo que el algoritmo se divide en dos partes: obtención de los puntos característicos y descripción de la región alrededor de cada punto de interés; La primera parte consiste en obtener los llamados *key points*, esto consiste en detectar las regiones de la imagen en las que se producen diferencias de gradientes significativas a ambos lados de dicho punto, es decir son pequeñas regiones con un nivel de intensidad estable. Para la descripción de cada punto característico el algoritmo realiza el cálculo de un histograma de direcciones del gradiente local alrededor de los puntos de interés [Alegre y Fernández, 2016].

Bag of visual visual words

El modelo *Bag of Visual Words* (BoVW) basado en el modelo *Bag of words* nos permite representar a la imagen mediante la construcción de un vocabulario de palabras visuales; esto lo realiza utilizando las características locales definidas por algún descriptor como SIFT [Valveny, 2018]. Para crear un diccionario es importante

contar con un conjunto de imágenes pertenecientes a una misma clase (*dataset*) de la que se obtienen los puntos de interés (*Keypoints*), generando el vector de características, estos vectores se agrupan entre ellos mediante alguna técnica de *clustering* (K Means, *Gaussian Mixture Model*) para encontrar las palabras visuales representativas a cada grupo de características similares. Al contar con el diccionario cada imagen del conjunto (*dataset*) se representa mediante un histograma que contiene con qué frecuencia aparece en esa imagen cada una de las palabras visuales del diccionario [González *et al.*, 2016]

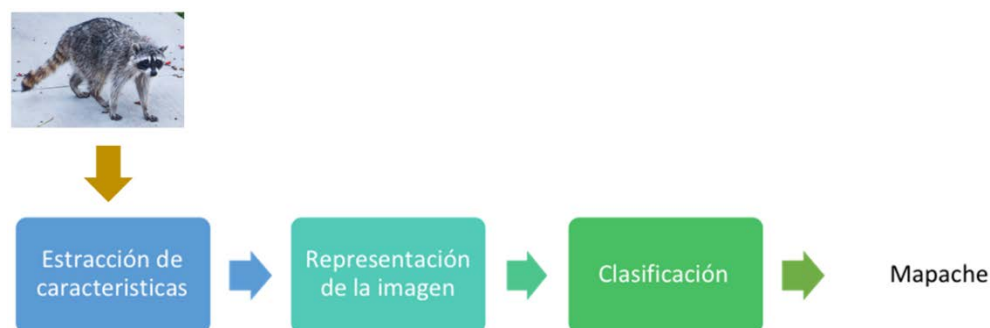
Support vector machines

Una máquina de vectores de soporte (SVM por sus siglas en inglés) es un clasificador discriminativo que se define utilizando un hiperplano de separación entre las clases, siendo el hiperplano la versión de una línea N-dimensional, es utilizado al contar con datos de entrenamiento etiquetados (aprendizaje supervisado) y un problema de clasificación binaria. Originalmente SVM era una técnica para construir un clasificador de dos clases, pero actualmente puede ser utilizado en problemas de regresión y agrupamiento. Si fuera el caso de tres dimensiones se plantea el uso de un *kernel*. Se mapea el vector de características en un espacio de dimensiones superiores utilizando la función de *kernel* y construyendo una función de discriminación lineal óptima en el espacio o hiperplano óptimo que se ajusta a los datos de entrenamiento. Una vez que se cuenta con los hiperplanos, se puede utilizar este modelo matemático para clasificar datos desconocidos en función de donde se encuentre en este plano.

2. Métodos

Para realizar la clasificación de las imágenes se realizó una aplicación que siguió un esquema basado en tres etapas las cuáles se observan en la figura 1. La primera etapa abarca la extracción de características de la imagen a clasificar, en la siguiente etapa se realiza la representación de la imagen a partir de la lista de puntos claves obtenidos por el algoritmo de extracción de características SIFT y por último se realiza la clasificación de la imagen mediante la generación de una etiqueta. En

cada una de las etapas se trabajó con el lenguaje de programación de Python y la librería para la manipulación de imágenes OpenCV.



Fuente: Basado en [Valveny, 2018].

Figura 1 Esquema de clasificación de imágenes.

Extracción de características

Previo al comienzo de la extracción de características se definió a las especies búho, mapache y pez como las tres primeras clases a identificar por el programa clasificador, para lo cual se realizó un concentrado de imágenes representativas de dos especies de la sierra Nor Oriental de Puebla y una tercera clase pez que engloba de manera general imágenes de peces, el banco de imágenes se generó de muestras aleatorias obtenidas del sitio de internet image-net.org que permite la manipulación de las imágenes bajo la licencia de uso de investigación o de educación. El objetivo de clasificar es determinar de forma automática el contenido de las imágenes capturadas en campo por una foto trampa, permitiendo identificar a qué clase pertenece; para lograr esto se comienza con la extracción de las características, que consiste en comprender el contenido de la imagen encontrando las regiones de interés a las que llamaremos puntos clave.

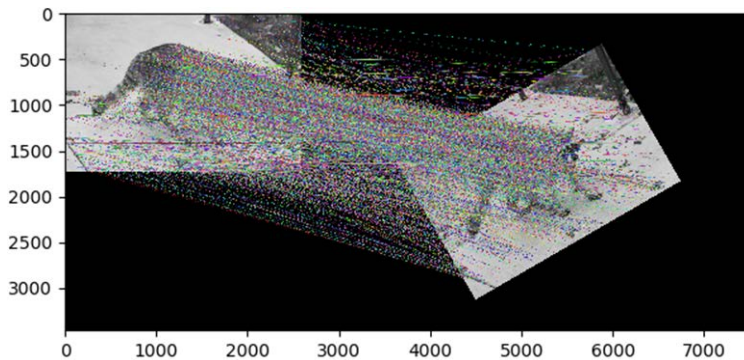
La caracterización del objeto que deseamos clasificar se realiza buscando puntos clave que se generen de forma invariable a la escala, posición o contraste de la imagen que estamos identificando, los puntos clave se obtienen utilizando el protocolo SIFT (*Scale Invariant Feature Protector* por sus siglas en inglés), éste algoritmo basa su funcionamiento en la diferencia de filtros gaussianos a diferentes resoluciones de la imagen analizada, permitiendo buscar puntos que representan pequeñas regiones con un nivel de intensidad estable, en el que se produce una

variación significativa del gradiente en diferentes direcciones; para la detección de los puntos clave se utiliza la clase `xfeatures2d.SIFT` de la biblioteca OpenCV. La figura 2 muestra las pruebas en la obtención de los puntos clave (*key points*) de la clase mapache. Adicionalmente, para corroborar la efectividad en la invarianza en escala, posición y contraste del algoritmo SIFT se realizó una prueba en la que se modificó la orientación, brillo y tamaño de la imagen de la clase mapache, en la figura 3 se puede observar el resultado de la obtención de los puntos clave de las dos imágenes y la correlación entre los puntos mediante líneas que hacen resaltar la coincidencia de las dos imágenes.



Fuente: [<https://unsplash.com>].

Figura 2 Generación de puntos de interés.



Fuente: Elaboración propia.

Figura 3 Relación de puntos de interés.

Como se observa en las dos figuras anteriores la detección de los puntos clave permite identificar zonas relevantes de la imagen analizada. El algoritmo SIFT realiza dos actividades: la detección de los puntos de interés y la descripción de los mismos mediante la generación de un vector de características [Valveny, 2018] dicho vector se genera de 128 elementos, basados en la concatenación del

histogramas de gradientes, donde cada punto de interés es representado por una tupla de cuatro elementos (x, y, σ, θ) donde (x, y) representan las coordenadas del punto, σ es la escala obtenida y θ la orientación de dicho punto.

El vector de características se obtiene mediante la obtención de la magnitud y la orientación del gradiente en una ventana de 16x16 píxeles, se calcula un histograma de orientaciones del gradiente de 8 bins para cada una de las 16 regiones, por último, se normaliza el vector de 128 elementos [González, Alegre, Fidalgo, 2018]. En figura 4 se observa el vector de características generado por la aplicación.

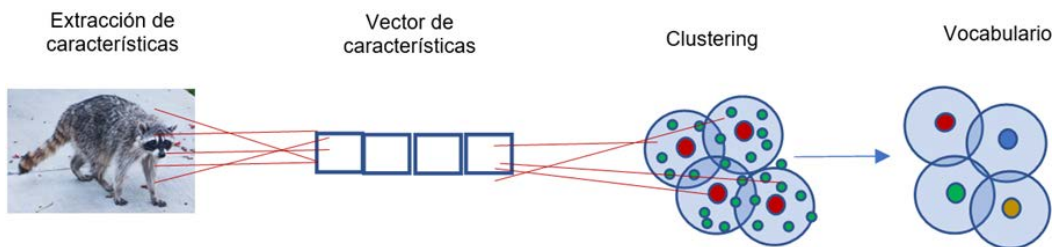
```
0.06779661, 0.00847458]], 'label': 'Mapache'}, {'feature_vector': array([[0.06779661, 0.00847458, 0.03225806, 0.01290323, 0.04516129, 0.05806452, 0.00645161, 0.01290323, 0.02580645, 0.01290323, 0.04516129, 0.00645161, 0.03225806, 0.01290323, 0.01935484, 0.0516129, 0.04516129, 0.00645161, 0.07096774, 0.02580645, 0.03225806, 0.03870968, 0.00645161, 0.18064516, 0.01290323, 0.01935484, 0.03225806, 0.01935484, 0.02580645, 0.01290323, 0.01290323, 0.06451613, 0.01935484]], 'label': 'Mapache']]
```

Fuente: Elaboración propia.

Figura 4 Vector de características.

Modelo bag of visual words

Después de contar con el vector de características se prosigue con la fase de representación de la imagen dado nuestro esquema de clasificación de imágenes, para esta etapa se utilizó el modelo de bag of visual words para crear un diccionario de palabras visuales mediante el proceso de agrupamiento (clustering), que ayude en la etapa de clasificación, la figura 5 muestra las etapas para la generación del diccionario.



Fuente: Elaboración propia.

Figura 5 Vector de características.

El diccionario es el resultado del uso del modelo Bag of words, en el que cada imagen es representada como un histograma de palabras visuales generado por la

detección y descripción de las características locales de las imágenes analizadas, esto es una forma de agregar todas las características locales en una única representación, el diccionario se generó después del análisis de 14 imágenes de las tres clases, cada imagen brindó información diferente según sus puntos clave.

Nuestro diccionario visual contiene los histogramas de los descriptores de características asociados a una etiqueta que se especificó al momento de realizar la extracción de puntos clave, definiendo tres etiquetas correspondientes a las tres clases con las que cuenta nuestro clasificador. La generación del diccionario se realizó a partir de los N vectores de características generados a partir del conjunto de imágenes que se utiliza para el entrenamiento

El clasificador se apoya en el uso del aprendizaje computacional, que consiste en definir una clasificación multiclase, para establecer múltiples fronteras del conjunto de imágenes representativas a las tres clases que se están analizando; este tipo de aprendizaje se constituye como supervisado al establecer una función basada en el etiquetado y el descriptor de los puntos de interés de las tres clases y que permite generar las tres categorías de búho, mapache y pez; para definir las fronteras entre estas tres clases, se establece el entrenamiento que genera el aprendizaje y que separa cada imagen en diferentes clases, para esta etapa se utilizó la biblioteca scikit-learn que permite hacer un aprendizaje automático para problemas supervisados, aprendiendo las propiedades de las diferentes clases, tomando información de las clases búho, mapache y pez de nuestra bolsa visual y comparando las propiedades de los conjuntos de datos, para poder predecir a que categoría pertenece y generar un estimador de clasificación.

3. Resultados

Para verificar la eficiencia de nuestro clasificador, se realizaron varias pruebas de etiquetado con imágenes diferentes a las utilizadas en la creación de nuestro diccionario visual de entrenamiento y etiquetado de nuestro clasificador, como métrica se tiene el resultado de utilizar la técnica de matriz de confusión, que es el estándar para identificar y resumir el rendimiento en la exactitud de nuestro clasificador al realizar la identificación o etiquetado de las imágenes muestra. La

matriz de confusión se emplea en análisis para aprendizaje supervisado ésta se genera de acuerdo con el número de clases que se desea evaluar, siendo ésta representada por una matriz de $n \times n$ donde las columnas representan el número de predicciones de cada clase, y la filas hacen alusión a las instancias de nuestras clases reales.

Para que el análisis nos arroje un resultado favorable es indispensable contar con un conjunto de datos diferentes al utilizado en el entrenamiento, realizar el número de pruebas iguales para cada clase n que se pretende clasificar. En nuestro caso se realizó una matriz de confusión multi clase como lo muestra la tabla 1.

Tabla 1 Matriz de confusión.

	Pez	Mapache	Búho
Pez			
Mapache			
Búho			

Fuente: Elaboración propia.

Se comenzó con el análisis por parte del clasificador de las clases búho, mapache y pez, establecidas al inicio de la investigación, la figura 6 da evidencia de cómo el clasificador realiza el etiquetado; cada clase cuenta con un total de catorce muestras obtenidas del sitio <https://unsplash.com/> que tiene una licencia de uso libre de las imágenes; los resultados del análisis de la clasificación de cada una de las imágenes se muestran en la tabla 2, como se puede observar los números de la diagonal representan los valores positivos reales, es decir aquellas pruebas en las que el clasificador etiquetó correctamente cada clase.

```
/Users/lespejo/miniconda3/envs/myenv/bin/py  
/Users/lespejo/miniconda3/envs/myenv/lib/py  
if diff:  
Output class: ['Buho']  
  
Process finished with exit code 0
```

Fuente: Elaboración propia.

Figura 6 Clasificación de una imagen.

Tabla 2 Matriz de confusión primer análisis.

	Pez	Mapache	Búho	Total de pruebas
Pez	0	0	14	14
Mapache	0	0	14	14
Búho	0	0	14	14

Fuente: Elaboración propia.

El número de pruebas para cada clase corresponde al total de la sumatoria de cada fila, en el caso de la clase mapache su total es de 14 y así sucesivamente para las otras clases.

La matriz multiclase que se ha generado arroja diversa información de clasificación como falsos negativos: que es la suma de los valores de cada fila excluyendo a los valores positivos, en el caso de la clase pez y mapache el valor fue de 14 lo que indica que fueron clasificados erróneamente y que pertenecen a una clase diferente a la etiquetada. Falsos positivos: es la sumatoria de cada una de las columnas excluyendo a los valores reales positivos, para las clases de pez y mapache fue de cero y en el caso de búho se tiene una sumatoria de 28 indicando que fueron clasificados como positivos y que pertenecen a la clase búho, cuando no corresponden a esa clase. Por último, se tiene a los reales negativos que se obtiene al sumar todas las columnas y filas excluyendo a la fila y columna de la clase que se desea analizar, para las clases pez y mapache el resultado fue de 28 y para la clase búho de 0 indicando que la única clase que no fue etiquetada de forma errónea es la de búho. A partir de la matriz de confusión podemos obtener las métricas de exactitud y precisión; para obtener la exactitud se deben sumar todos los resultados positivos entre el total de muestras realizadas, $exactitud = \frac{(TP_m + TP_p + TP_b)}{\sum_1^n P_T}$, para esta primera prueba de clasificación, la exactitud fue del .333; la precisión de cada una de las tres clases se calcula mediante la fórmula $precisión = \frac{TP}{TP + FP}$, para las clases de mapache y pez no existe precisión, dado que para estas dos clases los resultados de reales positivos y falsos positivos es de 0 y no se puede realizar la operación de división de entre 0, para el caso de la clase búho la precisión fue del 0.333 porque aunque todas las predicciones de sus imágenes fueron etiquetadas como búhos, tuvo falsos positivos de las otras clases.

Para validar o contrastar los resultados obtenidos de nuestra matriz de confusión en el primer análisis, se realizó nuevamente la extracción de características del conjunto de clases representativas, así como la generación del diccionario visual y el entrenamiento de nuestro clasificador de imágenes, al terminar el entrenamiento se realizó la prueba de clasificación de cada una de las imágenes de las tres clases, obteniendo los resultados mostrados en la segunda matriz de confusión (tabla 3).

Tabla 3 Segunda muestra de la matriz de confusión.

	Pez	Mapache	Búho	Total de pruebas
Pez	5	1	8	14
Mapache	2	1	11	14
Búho	1	1	12	14

Fuente: Elaboración propia.

Esta vez los resultados fueron diferentes a la primera prueba de clasificación; ya que para este análisis cada clase obtuvo valores reales positivos, la clase pez 5, la clase mapache 1 y la clase búho 12, complementando los resultados de la matriz de confusión se describen los datos que acompañan a la matriz, para el caso de los falsos negativos la clase pez tuvo 9, la clase mapache 13 y la clase búho 2; para los falsos positivos en la clase pez existieron 3, en la clase mapache 2 y en la clase búho 19 y los reales negativos obtenidos para la clase pez fueron 13, en la clase mapache 26 y por último en la clase búho 9.

La exactitud para esta nueva clasificación es de 0.428 mejorando la precisión por clase respecto al análisis anterior, los resultados obtenidos son clase pez 0.625, clase mapache 0.33 y 0.38 para la clase búho.

4. Discusión

El objetivo de la investigación es generar y contar con un clasificador que nos permita realizar la identificación de una especie a partir de las imágenes suministradas; hasta el momento el clasificador realiza un etiquetado con sesgos hacia una determinada clase, esto puede deberse a que en el momento de generar el diccionario de palabras visuales a partir de las N características extraídas de las imágenes de entrenamiento de cada clase, las muestras seleccionadas no son lo

suficientemente representativas para distinguir de una clase u otra al momento de realizar la clasificación. Aunque la clasificación se basa en un aprendizaje supervisado hay que tomar en cuenta que las imágenes seleccionadas tanto para el entrenamiento y grupo de pruebas generan un grupo de descriptores que no garantiza una clasificación óptima, además que se realizó una clasificación multi clase.

Si lo que desea es contar con un clasificador con una exactitud confiable, se hace indispensable trabajar desde la etapa de extracción de características utilizando algoritmos alternativos como ORB o SURF que nos permitan hacer un análisis y comparación de las características extraídas de las imágenes a partir de estos dos algoritmos, permitiendo así ayudar a reforzar el diccionario de palabras visuales y garantizar que al momento de realizar el entrenamiento las clases se encuentren plenamente definidas. Como el clasificador hace uso del entrenamiento supervisado, se necesita introducir un aprendizaje de refuerzo, al brindar retroalimentación a nuestro clasificador al momento que presente datos erróneos en el etiquetado de la especie analizada, la retroalimentación obliga a obtener las características de la imagen para incorporarlas al diccionario de datos además de ejecutar nuevamente el entrenamiento y reacomodar las clases con las características nuevas, para que en los posteriores análisis se incremente la exactitud y la confiabilidad global del clasificador.

5. Bibliografía y Referencias

- [1] Alegre E., Fernández L. (2016). Conceptos y Métodos en Visión por Computador. En Scale Invariant Feature Transform(131-140.). España: Grupo de visión del comité español de Automático.
- [2] Cohen, C. J. Haanpaa D., & Zott, J. P. (2015). Machine vision algorithms for robust animal species identification, IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, 2015, pp. 1-7. doi: 10.1109/AIPR.2015.7444526.
- [3] Cohen, C. J., Haanpaa, D., Rowe, S and Zott, J. P. (2011). "Vision algorithms for automated census of animals," 2011 IEEE Applied Imagery Pattern

- Recognition Workshop (AIPR), Washington, DC, pp. 1-5. doi: 10.1109/AIPR.2011.6176371.
- [4] Cámara de Diputados del H. Congreso de la Unión. (2018). Ley General de Vida Silvestre. Ciudad de México: http://www.diputados.gob.mx/LeyesBiblio/pdf/146_190118.pdf.
- [5] CONAFOR (2017). Informe final del proyecto denominado establecimiento de estaciones de monitoreo de diversidad de la fauna silvestre en predios bajo manejo forestal incorporados a la Asociación Regional de Silvicultores de Pátzcuaro- Tierra Caliente AC: https://www.conafor.gob.mx/EstudiosRegionalesForestales/ERF_Estudio_Fauna_UMAFOR_1608.pdf
- [6] Csurka, G., C. R. Dance, L. Fan, J. Willamowski, & C. Bray. Visual Categorization with Bags of Keypoints. Workshop on Statistical Learning in Computer Vision. ECCV 1 (1–22), 1–2.
- [7] Duarte M. M L. Fernández, Chang. Clasificación de Objetos en Imágenes Usando SIFT, 2015 septiembre.
- [8] Gallina, S, López, C. (2011). Manual de Técnicas para el estudio de la Fauna. Querétaro: [https:// https://bit.ly/2OE2OFg](https://bit.ly/2OE2OFg).
- [9] Garrido, G., Josh, P. (2018). OpenCV 3.x with Python By Example Second Edition. Birmingham UK: Packt Publishing Ltd.
- [10] González, V., Alegre, E., & Fidalgo, E. (2016). Clasificación de imágenes con Bag of Visual Words. En *Conceptos y Métodos en Visión por Computador* (183-200). España: Grupo de visión del comité español de Automático.
- [11] Hernández, B., Sánchez, A., & Delfín, C., Ochara'n, O., Morteo, E., Ríos, E. (2018, mayo 15). Desarrollo de aplicación para el conteo automático de murciélagos en cuevas basado en visión por computadora. *Research in Computing Science*, 147(7), pp. 11-22. 2019, agosto 20, De DBLP metadata for RCS Base de datos.
- [12] Lowe, D.G. *International Journal of Computer Vision* (2004) 60: 91. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [13] Rico, A., Rabuñal, J., Agudo, J., Pena, L., & Rodriguez, A. (2015). Un sistema de detección de peces para escala de hendidura vertical utilizando tecnología

láser y técnicas de visión artificial. Ingeniería del agua. 19. 229.
10.4995/ia.2015.3472

- [14] Sadawi, N. [jcensan]. (2014, 30 de agosto). Evaluating Classifiers: Confusion Matrix for Multiple Classes. [Video]: <https://www.youtube.com/watch?v=FAr2GmWNbT0>.
- [15] SEMARNAT (2015). Biodiversidad. agosto 20,2019, de Semarnat: https://apps1.semarnat.gob.mx:8443/dgeia/informe_resumen14/04_biodiversidad/4_1.html.
- [16] Sucar, E., Gómez G. (2010). Visión Computacional. Puebla, México: Instituto Nacional de Astrofísica, Óptica y Electrónica.
- [17] Valveny, E. [jcensan]. (2018, enero 20). [Video]: <https://bit.ly/374sX5u>.