



**HAL**  
open science

## Un Sistemas para Detección de Contaminación por Hidrocarburos: Aplicación al Oriente Ecuatoriano.

Jose Ignacio Pelaez, Gustavo Vaccaro, Estela R. Yanez, Mehdi Saqalli

### ► To cite this version:

Jose Ignacio Pelaez, Gustavo Vaccaro, Estela R. Yanez, Mehdi Saqalli. Un Sistemas para Detección de Contaminación por Hidrocarburos: Aplicación al Oriente Ecuatoriano.. RISCO: Revista Iberoamericana de Sistemas, Cibernética e Informática, 2016, pp.180-185. hal-01683552

**HAL Id: hal-01683552**

**<https://univ-tlse2.hal.science/hal-01683552>**

Submitted on 28 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Un Sistema para Detección de Contaminación por Hidrocarburos: Aplicación al Oriente Ecuatoriano

**JOSE IGNACIO PELAEZ**

Departamento de Lenguajes y Ciencias de la Computación  
Universidad de Málaga. España  
Investigador Proyecto PROMETEO  
Universidad de Guayaquil  
Guayaquil, Ecuador  
jipeláez@uma.es

**GUSTAVO VACCARO**

Universidad de Guayaquil  
Guayaquil, Ecuador,  
gustavo.vaccarow@fcmf.ug.edu.ec

**ESTELA R. YANEZ**

Grupo de Investigación SIBI  
Universidad de Guayaquil, Ecuador  
estela.yanezb@ug.edu.ec

**MEHDI SAQALLI**

Geode Laboratoire, Francia  
mehdi.saqalli@univ-tlse2.fr

## RESUMEN

La concienciación por el medio ambiente es cada vez más notable, y la sociedad en general exige que se lleve a cabo un control más riguroso sobre los posibles focos de contaminación. Los derrames de hidrocarburos de petróleo son una de las principales fuentes de contaminación de suelos y aguas ya que ocasionan perturbaciones en los ecosistemas al afectar su estructura y bioprocesos, afectando de manera directa a plantas, animales y humanos, y sobre todo a las poblaciones de microorganismos, los cuales representan una importante parte de los ecosistemas y son claves para los procesos biogeoquímicos. Muchos casos de contaminación por hidrocarburos se producen en zonas de alto valor ecológico y al mismo tiempo de difícil acceso, como por ejemplo en el Oriente Ecuatoriano. Disponer de sistemas automáticos análisis de imágenes que puedan ser incorporados en como parte de vehículos aéreos AUV, sería un elemento de gran importancia tanto para el control como de disuasión. En este trabajo se propone un modelo basado en análisis digital de imágenes para la detección automática de contaminación por petróleo crudo en suelos y aguas.

El sistema ha sido aplicado a la contaminación por crudo en el Oriente Ecuatoriano.

**Keywords:** Digital image analysis, environment, petroleum, contamination, oil spill.

## 1. INTRODUCCIÓN

El respeto y cuidado por el medioambiente ha pasado de ser algo excepcional a una exigencia de la sociedad en general. Antes de la década de los 70 los objetivos se centraban de manera limitada a la depuración del aire y de agua, mientras que el suelo no era considerado ya que se suponía que tenía una capacidad de autodepuración ilimitada. Sin embargo, a partir de la declaración de la “Carta Europea de Suelos” de 1972, la sensibilidad mundial comienza a cambiar.

Los derrames de hidrocarburos de petróleo son una de las principales fuentes de contaminación de suelos y aguas ya que ocasionan perturbaciones en los ecosistemas al afectar a su estructura y bioprocesos. En la figura 1 se muestra una balsa de acumulación de crudo, mientras que en la figura 2 se muestra una

canalización de crudo junto a población autóctona y animales de dichas comunidades.

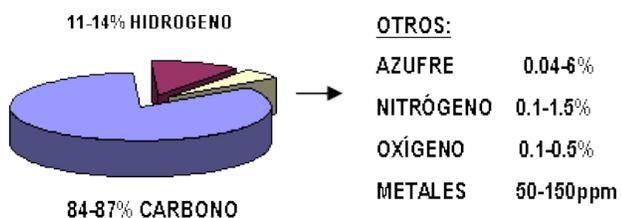


**Figura 1.** Balsa de acumulación de crudo.



**Figura 2.** Canalizaciones de crudo junto a población autóctona y animales.

Los hidrocarburos son compuestos de gran abundancia en la naturaleza integrados por átomos de carbono e hidrógeno, los cuales se clasifican según la estructura de los enlaces existentes entre los átomos de carbono que componen la molécula. En la figura 3 se muestra la composición del petróleo.



**Figura 3.** Composición del petróleo.

Los hidrocarburos estipulan una actividad económica de primera importancia a nivel mundial ya que son los principales combustibles fósiles, además sirven de

materia prima para todo tipo de plásticos, ceras y lubricantes. Pero son estas formas de valor elevado económico (crudo y derivados) las responsables de graves problemas de contaminación en el medio natural, y de posible corrupción que dificulta su control. La figura 4, muestra un vertido de crudo en la zona del Oriente Ecuatoriano, producido por una de las principales empresas petroleras internacionales.



**Figura 4.** Contaminación por crudo en el Oriente Ecuatoriano.

Una forma de vigilar los posibles derrames de crudo, especialmente en zonas de alto valor ecológico y de difícil acceso, y al mismo tiempo de servir de efecto disuasorio, es la utilización de sistemas de análisis automático de imágenes incorporados en sistemas aéreos AUV. En la figura 5 se muestra un dron con un sistema de cámara.



**Figura 5.** Dron DJI Phantom 4© con sistema de cámara digital integrado.

En este trabajo se propone de un sistema de detección de crudo en suelos y aguas mediante procesamiento digital de imágenes. El trabajo ha sido organizado como sigue: en la segunda sección, se presenta el sistema propuesto; en la tercera sección, se aplica a imágenes de contaminación por crudo en el oriente ecuatoriano; finalmente, el trabajo muestra una sección de discusión y conclusiones.

### 3. MODELO PROPUESTO

Para el desarrollo de este trabajo se propone un clasificador por entrenamiento supervisado, cuyo desarrollo se establece en 5 fases: Captura, Pre-procesamiento, Segmentación, Extracción de características y finalmente un proceso de Clasificación basado en redes neuronales. Un esquema del modelo propuesto se muestra en la Figura 6, donde asimismo se listan los componentes principales de cada fase. Estas serán detalladas a continuación.



**Figura 6.** Fases del modelo para categorización de contaminación petrolera.

#### Fase 1. Captura.

En la primera fase, las imágenes del terreno con posibilidad de contaminación son capturadas. Para ello se ha utilizado un dron DJI Phantom 4®, equipado con una cámara con resolución de 4K [1]. En este proceso se ha establecido una distancia al objeto máxima de 5 metros, con el propósito de mantener un detalle suficiente en las imágenes. En la figura 7 se muestran dos fotografías que han sido capturadas de esta manera, en las cuales se puede apreciar con claridad la presencia de contaminación por vertido de crudo.

#### Fase 2. Preprocesamiento

Considerando que las fotografías recogidas posean la menor cantidad de irregularidades producto del movimiento, pero sin afectar el rendimiento y velocidad de procesamiento del sistema, se ha empleado un filtro de medias móviles para suavizar

las imágenes capturadas [2]. Por otra parte, para reducir la complejidad del proceso, las dimensiones de las imágenes son reducidas a 800 x 450 píxeles.



(a)



(b)

**Figura 7.** Imágenes capturadas por dron donde se aprecia muestra de suelo, agua y vegetación contaminada con crudo de los pozos del Oriente Ecuatoriano.

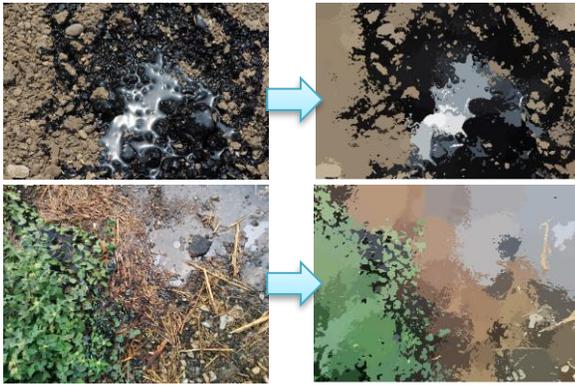
#### Fase 3. Segmentación

El propósito de la segmentación es descomponer la imagen en regiones con características de color similares, con la intención de identificar y delimitar objetos. En este caso, la preservación de bordes es crucial, por lo cual se ha utilizado un algoritmo de segmentación basado en Mean Shift [3]. El Mean Shift se fundamenta en una estimación de gradiente de densidad no paramétrico, no asume una morfología predefinida en la imagen, y ha sido probado como un método altamente eficiente para la segmentación automática de imágenes con preservación de bordes. [4], [5]. Matemáticamente, dado que  $g(x)$  sea un kernel gaussiano radialmente simétrico,  $h$  sea un parámetro controlado por el usuario que define el radio de una ventana, y  $x_i, i=1, \dots, n$  un set de puntos

en el espacio de dimensión  $d$  ( $\mathbb{R}^d$ , donde los puntos son píxeles de la imagen); para cada punto  $x_k$  ( $1 \leq k \leq n$ ), la secuencia de ubicaciones sucesivas del kernel  $y_j, j=0,1,2,\dots$  se define como:

$$y_{j+1} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{y_j - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{y_j - x_i}{h}\right\|^2\right)} \quad (2)$$

El Mean Shift ejecuta un procedimiento de gradiente ascendente sobre la función de probabilidad resultante; por lo que  $y_j$  converge a la moda de la función. Esto resulta en una agrupación de los píxeles en la imagen cuya función de probabilidad converge hacia la misma moda, y finalmente sustituyendo el color de los píxeles originales con el del representante del grupo. Un ejemplo de los resultados de la segmentación con Mean Shift se puede apreciar en la Figura 8.



**Figura 8.** Ejemplo de segmentación de imágenes con Mean Shift.

#### Fase 4. Extracción de Características

En esta fase se ha considerado un set de características amplio, consistente en tres tipos de modelos de descriptores cuantificables: en primer lugar se extrajeron descriptores de color a partir de los píxeles de la imagen; en segundo lugar, se extrajeron descriptores de color de los histogramas de la imagen; y finalmente, se extrajeron descriptores de textura. Todos los descriptores fueron extraídos de cada una de las componentes de los espacios de color RGB, HSI,  $L^*u^*v$ , y RGB normalizado.

Los descriptores de color de los píxeles de la imagen fueron procesados como la media y la varianza de cada componente de color. Las características extraídas de los histogramas de las imágenes de los fueron la posición y la altura de los dos picos más altos, la varianza, la asimetría, y la energía; cuya aplicación y validez ha sido comprobada en trabajos previos [6]–[8]. Finalmente, las características de

textura fueron evaluadas por medio de un análisis de entropía e los píxeles de la imagen [9]. Un total de 120 características fueron extraídas de esta manera.

#### Fase 5. Entrenamiento supervisado

La solución de detección de contaminantes por hidrocarburos propuesta se ha diseñado a manera de un sistema de clasificación supervisada binaria, basada en redes neuronales artificiales (RNA) [10]. Para este trabajo se empleó una arquitectura de Perceptrón multicapa con 120 neuronas en la capa de Entrada, una neurona en la capa de Salida (binaria), y un número variable de neuronas “ $H$ ” en la capa Oculta. El proceso de clasificación fue establecido de manera que cada región segmentada sea clasificada como “contaminante” o “no contaminante”. Por lo tanto, fue necesario realizar una etiquetación manual de cada una de las regiones segmentadas en todas las imágenes obtenidas. Esta tarea de etiquetado fue realizada mediante una aplicación especialmente diseñada que facilita la visualización de las regiones segmentadas y permite el etiquetado de manera sencilla. Estas etiquetas constituyen las salidas deseadas durante el proceso de entrenamiento del clasificador por redes neuronales.

El mejor número de neuronas en la capa oculta se computó al determinar el rendimiento resultante al entrenar un set de 100 RNAs, considerando en cada una de estas un incremento de 1 en el número de  $H$ , empezando en 20 y terminando en 120 neuronas ( $20 \leq H \leq 120$ ). Para esto, el conjunto total de regiones segmentadas en todas las imágenes fue dividido aleatoriamente en 3 grupos por cada ejecución de entrenamiento: 40% para el grupo de entrenamiento, 30% para el grupo de validación, y 30% para el grupo de pruebas.

El rendimiento del clasificador entrenado fue evaluado mediante el coeficiente de correlación de Matthews (MCC) [11], considerado como una medida de la calidad de la clasificación binaria, donde un  $MCC = 1$  representa una predicción perfecta, y un  $MCC = -1$  representa un completo desacuerdo entre la predicción y la observación. El MCC ha sido computado como:

$$MCC = \frac{VP \times VN - FP \times FN}{\sqrt{(VP + FP) \times (VP + FN) \times (VN + FP) \times (VN + FN)}} \quad (3)$$

Donde  $VP$  representa el número de verdaderos positivos,  $VN$  el número de verdaderos negativos,  $FP$  el número de falsos positivos, y  $FN$  el número de falsos negativos. Finalmente, el clasificador que alcanzó el mejor desempeño fue seleccionado a partir de su MCC asociado.

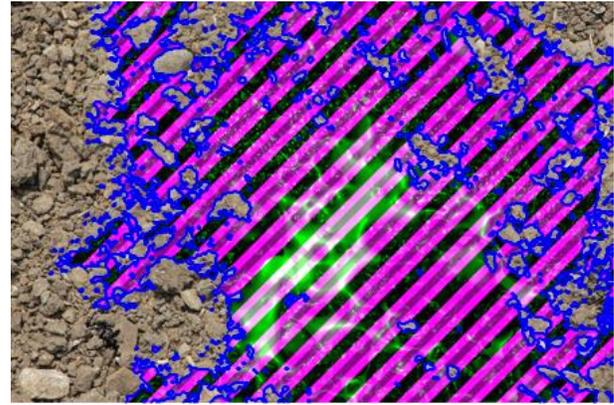
#### 4. RESULTADOS

En la fase de captura, se ha obtenido un total de 250 imágenes de diferentes terrenos y con distintos niveles de contaminación por hidrocarburos. Posteriormente, estas imágenes han sido pre-procesadas para disminuir los efectos del movimiento, y redimensionadas. En la fase de segmentación se consiguió identificar un promedio de 332 regiones por imagen, dando un total de 83015 regiones individuales. Estas regiones fueron etiquetadas por un equipo de 3 expertos en contaminación ambiental provistos por la Universidad de Guayaquil como “contaminante” o “no contaminante”. A partir de estos datos de entradas y salidas se realizó el entrenamiento supervisado de un set de RNAs, y se determinó que la mejor configuración obtenida se compone de 86 neuronas en la capa oculta, obteniéndose un rendimiento de clasificación de  $MCC = 9.23$ . Un ejemplo de resultados de la identificación automática de contaminantes en imágenes digitales se aprecia en la Figura 9, donde las regiones clasificadas como contaminantes de manera automática por el sistema han sido marcadas con líneas diagonales, además de un filtro de superposición de color para ser identificadas de mejor manera por los usuarios.

#### 5. DISCUSION y CONCLUSIONES

En este trabajo se ha presentado un sistema para la detección de crudo en suelos y aguas mediante procesamiento digital de imágenes. Para ello se ha diseñado un modelo para el análisis de imágenes obtenidas por drones, donde finalmente se obtiene un clasificador entrenado de manera supervisada que posee la habilidad de identificar regiones en las imágenes que corresponden con contaminación originada por hidrocarburos. Cabe resaltar en esto la sencillez de la idea, ya que trata de obtener los objetos que componen la imagen de manera similar a como lo hace el cerebro humano; y la segunda es la complejidad de su implementación, debido a la dificultad de simular la clasificación de regiones pobremente delimitadas en entornos con características variables, tal como es el Oriente Ecuatoriano.

Uno de los componentes clave ha sido la identificación de los diferentes objetos que componen la imagen. Para ello se ha utilizado el algoritmo de Mean-Shift, el cual está siendo utilizado cada vez con más frecuencia en aplicaciones donde se requiere de precisión al momento de identificar bordes entre regiones.



(a)



(b)

**Figura 9.** Resultados de la identificación y automática de contaminantes por hidrocarburos en imágenes de suelo, agua y vegetación contaminada con crudo de los pozos del Oriente Ecuatoriano. Estas se corresponden con las imágenes mostradas en la Figura 7

El análisis realizado sobre la muestra obtenida permite concluir que es posible efectuar una identificación automática de regiones contaminadas por hidrocarburos sobre imágenes obtenidas por drones, de una manera altamente eficiente; sin necesidad de trasladar personal hacia puntos de difícil acceso. Por otra parte, el análisis de una gran cantidad de imágenes es una tarea altamente exigente para ser realizada por un escaso personal de expertos, por lo que el sistema propuesto permite optimizar los recursos existentes en este sentido.

Por otra parte, consideramos que en futuros trabajos en este ámbito podrían involucrar la generación de nuevos descriptores especializados para la identificación de hidrocarburos, y estudios de comparación de diversos modelos de caracterización.

## Agradecimientos

Esta investigación está financiada por la SENESCYT en el marco del proyecto PROMETEO del Gobierno de la Republica de Ecuador, y por el Proyecto MONOIL, dentro del convenio de cooperación Franco-Ecuatoriano del Institut de recherche pour le développement (IRD) de Francia y la Universidad de Guayaquil, Ecuador.

## 6. REFERENCIAS

- [1] “DJI, Shenzhen, China.” .
- [2] F. Veredas, H. Mesa, and L. Morente, “Binary tissue classification on wound images with neural networks and bayesian classifiers.,” *IEEE Trans. Med. Imaging*, vol. 29, no. 2, pp. 410–27, Feb. 2010.
- [3] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [4] H. Gonçalves, J. A. Gonçalves, and L. Corte-Real, “HAIRIS: a method for automatic image registration through histogram-based image segmentation.,” *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 776–89, Mar. 2011.
- [5] C. M. Christoudias, B. Georgescu, and P. Meer, “Synergism in low level vision,” in *Object recognition supported by user interaction for service robots*, 2002, vol. 4, pp. 150–155.
- [6] G. Vaccaro, J. I. Pelaez, and J. A. Gil, “Choosing the best image processing method for masticatory performance assessment when using two-coloured specimens.,” *J. Oral Rehabil.*, p. On Press, Mar. 2016.
- [7] J. I. Pelaez, G. Vaccaro, and A. Guijarro, “Un Modelo para la Categorización de Hormigones Mediante Procesamiento Digital de Imágenes,” in *19th World Multiconference on Systemics, Cybernetics - CISCI 2015*, 2015.
- [8] G. Vaccaro and J. I. Pelaez, “Dental tissue classification using computational intelligence and digital image analysis,” in *Biodental Engineering III - Proceedings of the 3rd International Conference on Biodental Engineering, BIODENTAL 2014*, 2014, pp. 221–226.
- [9] M. Nixon, *Feature Extraction & Image Processing*. London: Elsevier, 2008.
- [10] M. Jabarouti and H. Soltanian-Zadeh, *Artificial Neural Networks - Methodological Advances and Biomedical Applications*. InTech, 2011.
- [11] B. W. Matthews, “Comparison of the predicted and observed secondary structure of T4 phage lysozyme,” *Biochim. Biophys. Acta - Protein Struct.*, vol. 405, no. 2, pp. 442–451, Oct. 1975.