

**SEGMENTACIÓN SEMÁNTICA PARA IMÁGENES DE PAISAJES
TROPICALES**

GERMÁN DARÍO CORZO USSA

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE ELECTRÓNICA
MAESTRÍA EN INGENIERÍA ELECTRÓNICA
BOGOTÁ D.C.
2017**

**SEGMENTACIÓN SEMÁNTICA PARA IMÁGENES DE PAISAJES
TROPICALES**

GERMAN DARIO CORZO USSA

**Trabajo de grado presentado como requisito para optar al título de
MASTER EN INGENIERÍA ELECTRÓNICA**

**Director:
JULIÁN ARMANDO QUIROGA SEPÚLVEDA PhD.**

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE ELECTRÓNICA
MAESTRÍA EN INGENIERÍA ELECTRÓNICA
BOGOTÁ D.C.
2017**

TABLA DE CONTENIDO

1. INTRODUCCION	7
2. OBJETIVOS.....	9
2.1 OBJETIVO GENERAL.....	9
2.2 OBJETIVOS ESPECIFICOS.....	9
3. MARCO TEÓRICO.....	10
3.1 SEGMENTACIÓN.....	10
3.2 SEMÁNTICA.....	10
3.3 Descriptores de Covarianza.....	11
3.3.1 Obtención de puntos Sigma.	11
3.3.2 Algoritmo para presentación de datos	12
3.4 MÁQUINAS DE Vectores de soporte	12
3.4.1 SVM para MATLAB	14
3.5 Mezcla de Gaussianas para clasificación	14
4. ESPECIFICACIONES	16
4.1 PROCESO.....	16
4.2 ALGORITMO	16
4.3 REVISIÓN DEL ESTADO DEL ARTE.....	17
5. DESARROLLOS	20
5.1 CREACIÓN DE LA BASE DATOS.....	20
5.2 DEFINICIÓN DE CLASES	21
5.3 DEFINICIÓN DE LA SEMÁNTICA DE CLASIFICACIÓN	22
5.3.1 Intervención del conocimiento experto.	23
5.4 ALGORITMO DE CLASIFICACIÓN	25
5.4.1 Selección de parámetros de prueba	25
5.4.2 Selección de los datos de entrenamiento	27
5.4.3 Error de muestras para entrenamiento	28
5.4.4 Pruebas en imagen de entrenamiento.....	29
5.4.5 Pruebas de imagen segmentada con algoritmo por pixel	32
5.4.6 Pruebas y cálculo de error en imagen entrenada con datos de otra ..	32
5.4.7 Pruebas y cálculo de error en imagen clasificada con GMM de otra ..	33

5.5	Producto final del algoritmo de segmentación semántica.....	34
5.5.1	Pruebas y cálculo de error del método propuesto caso I	35
5.5.2	Pruebas y cálculo de error del método propuesto caso II	37
6.	ANÁLISIS DE RESULTADOS	39
6.1	análisis de error para el algoritmo de segmentacion	39
6.2	análisis de error para el método propuesto	40
6.2.1	Análisis de resultados por clase	41
6.2.2	Análisis de resultados casos de segmentación	43
7.	CONCLUSIONES	46
8.	RECOMENDACIONES Y TRABAJOS FUTUROS	47
9.	REFERENCIAS BIBLIOGRÁFICAS	48
10.	ANEXOS	50

LISTA DE TABLAS

Tabla 1. Características de toolbox Matlab para clasificadores SVM	14
Tabla 2. Revisión del estado del arte para segmentación.....	18
Tabla 3. Fragmento de la base de datos de imágenes tomadas con sensores CCD e IR.....	20
Tabla 4. Error de clasificación con 750 muestras, imagen 012M.tif	27
Tabla 5. Error de muestras (cuadros) para entrenamiento	28
Tabla 6. Matriz de confusión para cuadros de entrenamiento imagen 012M.tif.....	29
Tabla 4. Matriz de confusión para imagen 012M.tif clasificada con 50 muestras de ella misma	31
Tabla 5. Matriz de confusión para imagen 012M.tif clasificada con GMM.....	32
Tabla 6. Matriz de confusión para imagen 012M.tif clasificada con 50 muestras de la imagen 011M.tif.....	33
Tabla 7. Matriz de confusión para imagen 012M.tif clasificada GMM de la imagen 011M.tif.....	34
Tabla 8. Matriz de confusión imagen 011M.tif segmentada con 450 muestras de otras imágenes.....	36
Tabla 9. Matriz de confusión imagen 011M.tif mejorada con muestras de sí misma.....	36
Tabla 10. Matriz de confusión imagen 012M.tif segmentada 960 muestras de 39 imágenes	37
Tabla 11. Matriz de confusión imagen 012M.tif, mejorada con muestras de sí misma.....	38
Tabla 12. Píxeles correctamente clasificados en 5 imágenes evaluadas con información de sí mismas	39
Tabla 13. Píxeles correctamente clasificados en 5 imágenes con información de otra imagen similar.....	41

LISTA DE FIGURAS

Figura 1. GMM Modelo Mezcla Gaussiana a partir de dos componentes.....	15
Figura 2. Diagrama de bloques del proceso de segmentación.	16
Figura 3. Diagrama de flujo del algoritmo de segmentación semántica.	17
Figura 4. Características de las clases.	21
Figura 5. Imágenes de prueba segmentadas por expertos.	24
Figura 6. Sigma points en el espacio euclidiano.	26
Figura 7. Entrenamiento SVM toolbox Matlab con 750 muestras, diferente kernel	26
Figura 8. Comparación del error de clasificación de la imagen 012M.tif.....	27
Figura 9. Selección de 10 muestras por clase para entrenamiento de la imagen 012M.tif ..	28
Figura 10. Selección de cuadros de 10 x 10 para clasificación de la imagen 012M.tif	30
Figura 11. Imágenes de prueba clasificadas con muestras por clase y su imagen de error. 30	
Figura 12. Resultado imagen 012M.tif segmentada con algoritmo no supervisado de mezcla de gaussianas.....	32
Figura 13. Resultado imagen 012M.tif segmentada con 50 muestras de imagen 011M.tif .	33
Figura 14. Resultado imagen 012M.tif segmentada con GMM de imagen 011M.tif	34
Figura 15. Resultado imagen 0000M.bmp segmentada con algoritmo semántico.....	35
Figura 16. Resultado imagen 011M.tif, segmentada con método propuesto	35
Figura 17. Resultado imagen 011M.tif, mejorada con muestras de si misma	36
Figura 18. Resultado imagen 012M.tif, segmentada 960 muestras de 39 imágenes	37
Figura 19. Resultado imagen 012M.tif, mejorada con muestras de sí misma	37
Figura 20. Pixeles clasificados correctamente vs algoritmo de segmentación aplicado.....	41
Figura 21. Pixeles clasificados correctamente vs algoritmo de segmentación	42
Figura 22. Análisis de implementación del método adicionando datos de aprendizaje.....	44
Figura 23. Análisis de implementación del método adicionando datos de aprendizaje.....	45

1. INTRODUCCIÓN

Las imágenes cartográficas tienen un gran impacto en la seguridad nacional gracias a la información que pueden suministrar, pues dependiendo del tipo de sensor con el que sea generada es posible conocer accidentes topográficos, clima, vegetación, hidrografía, etc., que han servido enormemente a las fuerzas armadas para desarrollar su función constitucional en la defensa del estado. Las Fuerzas Militares cuentan con aeronaves que tienen sensores de dispositivo de carga acoplada (*en inglés charge-coupled device, conocido también como CCD*) para toma de imágenes diurnas y sensores térmicos infrarrojos (*IR*), los cuales producen imágenes de muy buena calidad en áreas geográficas muy amplias.

En los últimos cinco años han surgido necesidades para determinar a través de imágenes la detección de explotación ilegal de recursos naturales, el impacto sobre las fuentes hídricas y la conservación del medio ambiente. Las herramientas y recursos para hacer un análisis de mayor profundidad de las imágenes obtenidas son muy pocas o requiere que se actualicen o compren otros elementos que no están desarrollados para las aplicaciones específicas que en Colombia son relevantes. La detección de objetos generalmente se realiza de forma que sea distinguible a simple vista, lo cual resulta muy pobre al momento de tomar decisiones, además, los métodos computacionales para descubrir información detrás de las formas, texturas, color, bordes, contraste y brillo en las imágenes obtenidas son escasos, debido a que no es un recurso de fácil acceso y porque no había surgido la necesidad hasta ahora de analizar las imágenes para encontrar otro tipo de objetos de interés.

El análisis de imágenes para producción de cartografía es demasiado lento, poco eficiente y mucho menos confiable, ya que depende de una valoración subjetiva, además es muy difícil de actualizar al menos en periodos que permitan tomar decisiones. Se requiere una cantidad considerable de personal especializado, equipo y tiempo para procesar imágenes en un modelo fotogramétrico. En este sentido es necesario emplear una técnica, método o sistema que permita obtener más y mejor información en las imágenes a partir de aquellos objetos que por su relevancia permitan identificar patrones anormales, en aras de prever escenarios en los cuales se vea afectada la población, el medio ambiente y/o el desarrollo socioeconómico del país.

Debido a la dificultad de procesar grandes volúmenes de información por las condiciones donde se requiere obtener datos de relevancia, se debe obtener una herramienta versátil y de mínima carga computacional, que permita tomar decisiones de forma rápida, pero con alta probabilidad de ser cierta para no caer en el gasto de recursos innecesariamente.

Una técnica para lograr el entendimiento del entorno y la identificación de objetos, consiste en entrenar el sistema computacional a través de un conjunto de datos. Se propone un método empírico que reúna pruebas sistemáticas para capturar una gran cantidad de información en paisajes tropicales cubriendo la mayor cantidad de elementos, situaciones, entorno, clima, temporada luminosidad, entre otros y generando grupos con clasificadores que reúnan la mayor cantidad de muestras sobre los objetos de interés llamados clases, con el fin de lograr el etiquetado semántico de los pixeles con una amplia variedad de datos de entrenamiento. Con estos conjuntos de datos se permite la mejor comprensión de la escena y adaptarse a las necesidades futuras de los paisajes tropicales colombianos.

El conocimiento experto en este tipo de soluciones es fundamental para reducir la probabilidad de error, sin embargo, la segmentación manual de una imagen entera resulta muy costosa y poco eficiente, por esta razón el algoritmo propuesto parte de tomar las muestras de cada una de las clases a partir de una imagen elegida por el experto, la cual contenga suficientes ejemplos claros y que puedan ser elegidos por el analista de imágenes. De las demás imágenes a ser procesadas, solo se tomarán un pequeño número de muestras que permita observar ejemplos zonas de difícil decisión. Todas estas muestras alimentan una tabla donde se encuentran los datos etiquetados en 3 clases y que son entrenados a través de máquinas de vectores de soporte (*support vector machine* por sus siglas en inglés SVM), el programa se encarga de clasificar la imagen a través de una malla de procesamiento y se filtra el resultado para obtener los bordes iniciales. Al final la imagen es procesada con algoritmos para encontrar objetos “*extraños*” dentro de un paisaje tropical colombiano y se presenta una imagen completamente segmentada en sus clases.

Finalmente, en este trabajo se evalúa el algoritmo desarrollado a través del estudio del error con imágenes segmentadas cuidadosamente con ayuda del analista experto en imágenes cartográficas en las cuales servirán como plantilla de comparación. Este análisis se realiza encontrando pixel a pixel el error de clasificación y haciendo uso de la matriz de confusión, con base en los requerimientos del usuario el objetivo es obtener imágenes segmentadas correctamente mínimo con el 80% de certeza en la clasificación del total de pixeles por imagen.

2. OBJETIVOS

2.1 OBJETIVO GENERAL

Desarrollar un algoritmo de segmentación semántica a partir de imágenes obtenidas con sensores ópticos e IR capturadas desde una vista superior, para el análisis de paisajes tropicales colombianos.

2.2 OBJETIVOS ESPECIFICOS

- Crear una base de datos con imágenes clasificadas de paisajes tropicales obtenidos con sensores CCD e IR desde plataformas aéreas para el desarrollo y evaluación del algoritmo de segmentación.
- Revisar el estado del arte sobre técnicas de segmentación en color y textura, así como los las técnicas y métodos utilizados para procesamiento de imágenes cartográficas y paisajes forestales.
- Diseñar e implementar un algoritmo de segmentación para imágenes de paisajes tropicales.
- Definir un protocolo de evaluación para el algoritmo de segmentación desarrollado utilizando k-means o mezcla de Gaussianas como método de comparación.

3. MARCO TEÓRICO

3.1 SEGMENTACIÓN

La segmentación es el proceso de dividir una imagen digital en varias partes u objetos. El objetivo es simplificar y/o cambiar la representación de una imagen en otra más significativa y más fácil de analizar. Se usa tanto para localizar objetos como para encontrar los límites de estos dentro de una imagen. La segmentación se logra mediante el proceso de asignación de una etiqueta a cada píxel de la imagen de forma que los píxeles que compartan la misma etiqueta también tendrán ciertas características visuales similares; el resultado es un conjunto de segmentos que cubren en conjunto a toda la imagen. Cada uno de los píxeles de una región es similar en alguna característica, como el color, la intensidad o la textura. Regiones adyacentes son significativamente diferentes con respecto a uno o varios tipos de características.

Los algoritmos de segmentación se basan en los siguientes principios:

- Discontinuidades del nivel de gris: Consisten en segmentar la imagen a partir de los cambios grandes en los niveles de gris entre los píxeles. Las técnicas que utilizan las discontinuidades como base son la detección de líneas, de bordes, de puntos aislados.
- Similitud de niveles de gris: Es lo contrario al método anterior, las divisiones de la imagen se hacen agrupando los píxeles que tienen unas características similares. Algunas técnicas que usan esto son la umbralización, el crecimiento de regiones, etc.

3.2 SEMÁNTICA

El problema de segmentación semántica consiste en otorgar una etiqueta o categoría a cada píxel de una imagen. A diferencia de los sistemas de detección y reconocimiento de objetos, que dan como resultado la ventana rectangular donde se ha detectado un objeto y son evaluados por la precisión de estas ventanas, los sistemas de segmentación semántica tienen como objetivo delimitar precisamente los objetos de las distintas categorías a nivel de píxel, dando como resultado cualquier forma arbitraria [14].

El proceso de segmentación es aquel que divide una imagen en regiones cuyos píxeles poseen atributos similares. Idealmente cada región segmentada corresponde a un objeto que se encuentra en la escena. La segmentación se basa en el estudio de tres propiedades sobre las imágenes:

- Similitud: Cada uno de los píxeles de un elemento de la imagen tiene valores parecidos para alguna propiedad, por ejemplo sus componentes de color RGB.
- Discontinuidad: Los objetos se destacan del entorno y tienen, por tanto, unos bordes definidos, los cuales pueden ser arbitrarios o congruentes para lo cual pueden ser parte de polígonos regulares.
- Conectividad: Los píxeles pertenecientes al mismo objeto tienen que ser contiguos, es decir, deben estar agrupados, esta propiedad describe que los píxeles agrupados varían muy poco entre ellos.

El problema de la segmentación está dado en encontrar una partición en regiones homogéneas y sus fronteras. Se supone que las regiones homogéneas son objetos perceptualmente significativos y los bordes sus límites.

3.3 DESCRIPTORES DE COVARIANZA

Los descriptores de covarianza proveen una representación de características de baja dimensión que puede integrar de forma simple múltiples canales, como color, respuesta a filtros, bordes, entre otros y utiliza la correlación entre ellos. Las matrices de covarianza pueden ser eficientemente calculadas y proveen una integración de varios canales de características, sin embargo, el espacio de las matrices de covarianza no tiene la forma de espacio vectorial Euclidiano, esta representación no puede ser directamente usada para muchas técnicas de aprendizaje de máquina [13].

Para resolver este inconveniente y aprovechar el potencial de los descriptores de covarianza, se propone una técnica que realiza la transformación del espacio vectorial.

Para resolver este inconveniente y aprovechar el potencial de los descriptores de covarianza, se propone una técnica que realiza la transformación del espacio vectorial.

3.3.1 Obtención de puntos Sigma.

Para representar matrices de covarianza individuales directamente en un espacio vectorial Euclidiano, es posible realizar una transformación que consiste en encontrar un pequeño conjunto de $2d + 1$ vectores específicos $s_i \in \mathbb{R}^d$ llamados puntos sigma [8], los cuales se construyen de la siguiente forma:

$$s_0 = \mu \quad s_i = \mu + \alpha(\sqrt{\Sigma})_i \quad s_{i+d} = \mu - \alpha(\sqrt{\Sigma})_i \quad (1)$$

3.3.2 Algoritmo para presentación de datos

Para la representación de características se propone un algoritmo basado en *Puntos Sigma* el cual requiere el vector de medias μ^k y la matriz de covarianza Σ^k , así [9]:

1. Realizar una regularización simple

$$\Sigma^k = \Sigma^k + \epsilon I \quad (2)$$

2. Calcular la raíz cuadrada de la matriz tal que

$$\Sigma^k = LL^T \quad (3)$$

3. Calcular s_i^k de acuerdo a (1)

4. Construir el conjunto de Puntos Sigma

$$S^k = (s_0^k, s_1^k, \dots, s_{2d}^k) \quad (4)$$

3.4 MÁQUINAS DE VECTORES DE SOPORTE

La transformación a un espacio vectorial Euclidiano permite emplear diversos métodos de clasificación de los puntos obtenidos. Para el desarrollo del entrenamiento de máquina se propone un método aplicable a las características de color de los paisajes tropicales, teniendo en cuenta que las imágenes son generalmente de alta resolución y tamaño, que las características de los objetos son diferenciables a vista de un experto con alta probabilidad de acierto y además que con pocos datos se obtengan resultados en la segmentación con baja probabilidad de error.

Las máquinas de vectores soporte (SVM, del inglés Support Vector Machines) han sido ampliamente utilizadas con éxito en el campo de la visión artificial gracias a su simplicidad, bajo coste computacional y sólidos fundamentos teóricos. Las SVMs pertenecen a la categoría de los clasificadores lineales, puesto que inducen separadores lineales o hiperplanos, ya sea en el espacio original de los ejemplos de entrada, si éstos son separables o cuasi-separables (ruido), o en un espacio

transformado (espacio de características), si los ejemplos no son separables linealmente en el espacio original. La búsqueda del hiperplano de separación en estos espacios transformados, normalmente de muy alta dimensión, se hará de forma implícita utilizando las denominadas funciones kernel [3].

Los hiperplanos de separación son buenos clasificadores cuando los ejemplos son perfectamente separables o cuasi-perfectamente separables, también es importante tener en cuenta que el proceso de búsqueda de los parámetros que definen dichos hiperplanos se puede hacer independientemente de la dimensionalidad del problema a resolver buscando espacios transformados de separación óptimos. A cada uno de estos espacios se le denomina espacio de características.

Debido a que se requiere la reducción de datos y simplificación del método, el procedimiento basado en SVM se adapta a los objetivos del algoritmo de segmentación. Una vez obtenidos los puntos sigma el proceso de aprendizaje de máquina se resume de la siguiente forma [3]:

Sea $\Phi: \mathbb{X} \rightarrow \mathcal{F}$ la función de transformación que hace corresponder cada vector de entrada x con un punto en el espacio de características \mathcal{F} , donde $\Phi(x) = [\phi_1(x), \dots, \phi_m(x)]$ y $\exists \phi_i(x), i = 1, \dots, m$, tal que $\phi_i(x)$ es una función no lineal la idea entonces es construir un hiperplano de separación lineal en este nuevo espacio. La frontera de decisión lineal obtenida en el espacio de características se transformará en una frontera de decisión no lineal en el espacio original de entradas en este contexto, la función de decisión en el espacio de características vendrá dada por:

$$D(x) = (\omega_1 \phi_1(x) + \dots + \omega_m \phi_m(x)) = \langle \omega, \phi(x) \rangle \quad (5)$$

Donde $\Phi: \mathbb{X} \rightarrow \mathcal{F}$

Por tanto, una función kernel puede sustituir convenientemente el producto escalar, así, dado el conjunto de funciones base, $\Phi = \{\phi_1(x), \dots, \phi_m(x)\}$ el problema a resolver sigue siendo encontrar el valor de los parámetros $\alpha_i^*, i = 1, \dots, n$, que optimiza el problema dual, pero expresado ahora como:

$$\text{máx} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (6)$$

$$\text{s. a.} \sum_{i=1}^n \alpha_i y_i = 0 \quad (7)$$







$$0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \quad (8)$$

Se está ya en disposición de describir la metodología necesaria para resolver un problema de clasificación de ejemplos no separables linealmente. Concretamente, la función de decisión, donde el valor de los parámetros $\alpha_i, i = 1, \dots, n$, se obtendrán como solución al problema de optimización cuadrática, conocidos el conjunto de ejemplos de entrenamiento $(x_i, y), i = 1, \dots, n$, el kernel K, y el parámetro de regularización C.

3.4.1 SVM para MATLAB

La documentación de Matlab muestra las características para cada tipo de Kernel SVM con la ventaja de poder exportar cada modelo en código C para realizar predicciones.

Tabla 1. Características de toolbox Matlab para clasificadores SVM

Classifier Type	Prediction Speed	Memory Usage	Interpretability	Model Flexibility
 Linear SVM	Binary: Fast Multiclass: Medium	Medium	Easy	Low Makes a simple linear separation between classes.
 Quadratic SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Medium
 Cubic SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Medium
 Fine Gaussian SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	High — decreases with kernel scale setting. Makes finely detailed distinctions between classes, with kernel scale set to $\sqrt{P}/4$.
 Medium Gaussian SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Medium Medium distinctions, with kernel scale set to \sqrt{P} .
 Coarse Gaussian SVM	Binary: Fast Multiclass: Slow	Binary: Medium Multiclass: Large	Hard	Low Makes coarse distinctions between classes, with kernel scale set to $\sqrt{P} \cdot 4$, where P is the number of predictors.

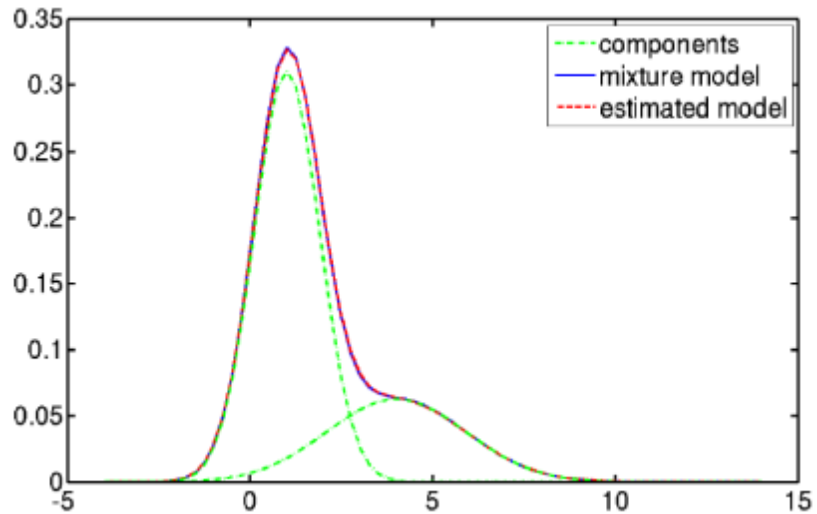
Fuente: MathWorks® Documentation, <https://www.mathworks.com/help/stats/choose-a-classifier.html>.

3.5 MEZCLA DE GAUSSIANAS PARA CLASIFICACIÓN

El modelo de mezcla de Gaussianas (GMM) se utiliza para representar la presencia de subpoblaciones dentro de una misma población y corresponde a una distribución que representa la distribución de probabilidad de alguna observación

en la población general. Son usados para crear inferencias estadísticas, aproximaciones y predicciones acerca de las propiedades de las sub poblaciones a partir de las observaciones o datos adquiridos de la población estudiada sin necesidad de información que identifique a la sub población.

Figura 1. GMM Modelo Mezcla Gaussiana a partir de dos componentes.



Fuente: http://catarina.udlap.mx/u_dl_a/tales/documentos/lmt/ramirez_a_e/capitulo3.pdf

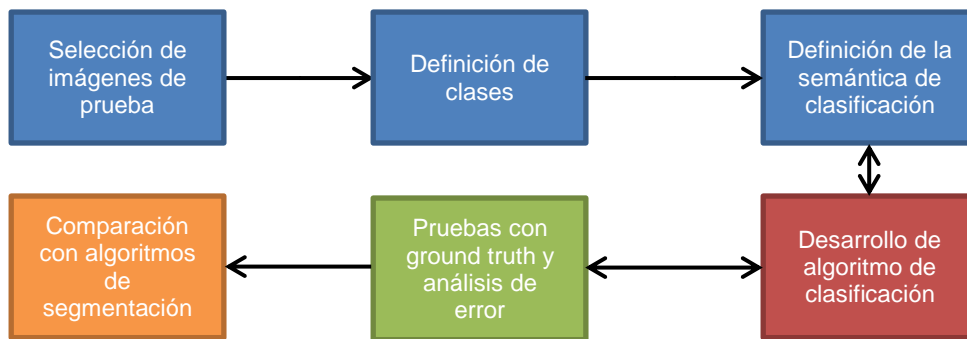
4. ESPECIFICACIONES

4.1 PROCESO

Se presenta un algoritmo para segmentación de imágenes que consiste en otorgar una etiqueta o categoría a grupos de píxeles en una imagen. La totalidad del algoritmo se implementó en el software Matlab. En primera instancia, al algoritmo de procesamiento se le presentan datos etiquetados sobre las imágenes, descritos por medio de los “sigma points”, a partir de los cuales se entrena una máquina de aprendizaje supervisado basada en SVM. A partir de esta máquina, se crea el modelo que es aplicado en el clasificador de “cuadros” de cada imagen que decide su clase a partir de los datos de entrada.

Finalmente, la segmentación semántica se presenta dividiendo la imagen de entrada en colores según la clase etiquetada. Los errores son calculados mediante la comparación píxel a píxel con una imagen de referencia (*ground-truth*) y con otra segmentada con un algoritmo de agrupación de píxeles, en este caso con un Modelo de Mezcla de Gaussianas (GMM). El procedimiento se desarrolla de acuerdo al diagrama de bloques de la figura 2.

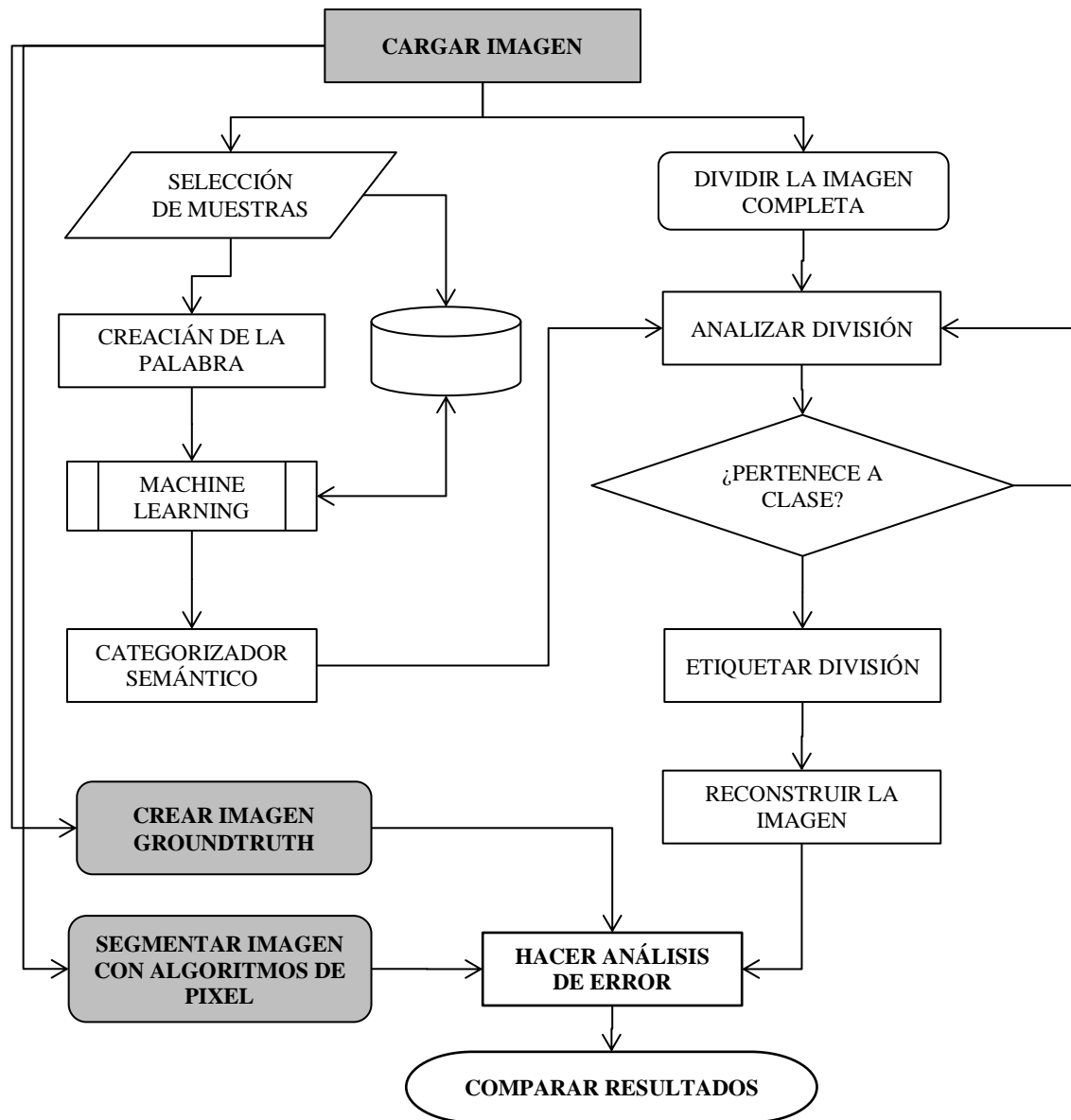
Figura 2. Diagrama de bloques del proceso de segmentación.



4.2 ALGORITMO

El diagrama de flujo para el desarrollo del algoritmo de segmentación se presenta en la figura 3.

Figura 3. Diagrama de flujo del algoritmo de segmentación semántica.



4.3 REVISIÓN DEL ESTADO DEL ARTE

Se hace una revisión del estado del arte sobre técnicas de segmentación en color y textura, así como los las técnicas y métodos utilizados para procesamiento de imágenes cartográficas y paisajes forestales.

Tabla 2. Revisión del estado del arte para segmentación

ARTÍCULO	DESCRIPCIÓN DEL MÉTODO
<p>Arbeláez, P.; Hariharan, B.; Gu, C.; Gupta, S.; Bourdev, L.; Malik, J.; University of California, Berkeley, <i>Semantic Segmentation using Regions and Parts</i></p>	<p>Se aborda el problema de segmentación y reconocimiento de objetos en imágenes del mundo real, se centra en categorías de objetos articulados como humanos y animales. Se propone un diseño para detectores de objetos basado en regiones que integra información descendente a partir de modelos de partes en ventanas de escaneo y señales de apariencia global. Los detectores producen puntajes específicos de clase para regiones ascendentes y luego agregan votos de múltiples candidatos superpuestos a través de la clasificación de píxeles. Se evalúa con respecto a la segmentación PASCAL (Página de objetos visuales VOC2011, disponible en: http://host.robots.ox.ac.uk/pascal/VOC/voc2011/index.html).</p>
<p>Avendaño, J.; Parra, J.; Bayona, J.; <i>Segmentación y clasificación de imágenes SAR en zonas de inundación en Colombia, una herramienta computacional para prevención de desastres. Junio 2014</i></p>	<p>Los datos de entrenamiento fueron generados con los resultados de las segmentaciones Fuzzy- Clustering, K-means y Region-Growing sobre imágenes SAR de zonas de inundación. Los métodos de clasificación implementados fueron un clasificador basado en el método Bayesiano y un clasificador basado en máquinas de vectores de soporte (SVM). Para evaluar el desempeño de los clasificadores se utilizaron índices como la exactitud total, la exactitud dependiendo del usuario, el índice Kappa y R'.</p>
<p>Castilla G.; <i>Tesis Doctoral: Object-oriented analysis of remote sensing images for land cover mapping: conceptual foundations and a segmentation method to derive a baseline partition for classification.</i></p>	<p>Producción de una partición con base en la estructura espacial de las imágenes en la que cada región puede considerarse relativamente homogénea y diferente de sus vecinas y que supere el tamaño de la unidad mínima cartografiable. Estas regiones son las unidades básicas de la clasificación, sobre las cuales se pueden definir una serie de atributos espaciales (forma, tamaño, orientación), estructurales (disposición interna, tono, textura y contraste entre las diferentes partes que las componen) y contextuales (relaciones con regiones vecinas) que no son aplicables a píxeles individuales.</p>
<p>Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B.; <i>The Cityscapes Dataset for Semantic Urban Scene Understanding</i></p>	<p>Se presenta Cityscapes, un paquete de referencia y un conjunto de datos a gran escala para entrenar y probar enfoques para el etiquetado semántico a nivel de ejemplos y a nivel de píxeles en el contexto de Deep Learning.</p>
<p>Girshick R.; Donahue J.; Darrell T.; Malik J.; "Rich feature hierarchies for accurate object detection and semantic segmentation", CVPR2014, provided by Computer Vision Foundations, 2014.</p>	<p>Se basa en dos ideas: (1) se pueden aplicar redes neuronales convolucionales (CNN) de alta capacidad a propuestas de la región ascendente para localizar y segmentar objetos y (2) cuando los datos de capacitación etiquetados son escasos, el preentrenamiento supervisado para una tarea auxiliar, seguida de un ajuste fino específico del dominio, produce un impulso de rendimiento significativo. El método es llamado Regiones con características de CNN. También se presentan experimentos que proporcionan información sobre lo que la red aprende, revelando la jerarquía de características de la imagen.</p>

<p>Kluckner S., Mauthner T., Roth P., Bischof H., Semantic classification in aerial imagery by integrating appearance and height information</p>	<p>Se presenta una técnica para obtener una clasificación semántica en el nivel de píxel capaz de integrar diversas modalidades, como el color, las respuestas de los bordes y la información de altura. Se propone una representación de características basada en cálculos de Sigma Points que permite una aplicación simple de descriptores de covarianza a un método de clasificación de árboles de decisión. Se incluye un conocimiento contextual semántico usando una formulación de campo aleatorio condicional. Finalmente se evalúa en la colección de imágenes de MSRC y luego se demuestra en tres conjuntos de imágenes aéreas: Dallas, Graz y San Francisco.</p>
<p>Herrero A.; "Algoritmos para la estimación de modelos de mezclas gaussianas (GMM)" Escuela técnica superior de Ingenieros Industriales y de Telecomunicación, Universidad de Cantabria, Julio – 2015.</p>	<p>Revisión para clasificación de datos almacenados para obtener información en el área del reconocimiento de patrones no supervisado. Se desarrolla en esta tesis el estudio de ciertos métodos de agrupamiento de datos y su aplicación a problemas concretos para imágenes, se considera principalmente modelo de mezclas Gaussianas (GMM) y otros conceptos relacionados como la Estimación de máxima verosimilitud y el algoritmo Kmeans.</p>
<p>Pacheco, C.; Del Valle, H.; Delrieux, C.; <i>Revisión de algoritmos de segmentación desde una perspectiva orientada al sensado remoto de imágenes ópticas y de radar, 2016.</i></p>	<p>Se analiza la literatura pasada y reciente sobre los principales algoritmos con aplicaciones en imágenes terrestres, tratando de encontrar procedimientos y protocolos computacionales que puedan usarse para segmentar cualquier tipo de imagen satelital, se abordan los siguientes algoritmos: Detección de bordes y regiones, Crecimiento de regiones, Híbrido entre crecimiento de regiones y clustering espectral, binarización con umbral local dinámico, segmentación contextualizada y encapsulada, colonia artificial de abejas, Level Set y Bayesiano</p>

5. DESARROLLOS

5.1 CREACIÓN DE LA BASE DATOS

Para la selección de muestras que permitieran hacer la definición y caracterización de las clases se realizaron pruebas sobre 5 imágenes de paisajes tropicales seleccionadas de un conjunto de 219 fotografías tomadas desde una plataforma aérea con sensores CCD e IR. La selección se hizo teniendo en cuenta la presencia de las clases de interés requeridas en cada una de ellas y verificando que fuesen homogéneas entre sí, es decir, tomadas en lo posible de forma vertical, a la misma hora del día y en condiciones de sol/nubosidad similares.

Tabla 3. Fragmento de la base de datos de imágenes tomadas con sensores CCD e IR

Nombre	Formato	Sensor	Vegetación	Agua	Construcciones	Otros	Reconocimiento visual
000MO	BMP	CCD	1	1	1	1	BUENA
001M	TIFF	CCD	1	1	1	1	BUENA
002M	TIFF	CCD	1	1	0	1	BUENA
003M	TIFF	CCD	1	1	0	1	MALA
004M	TIFF	CCD	1	1	0	1	REGULAR
005M	TIFF	CCD	1	1	0	1	MALA
006M	TIFF	CCD	1	1	1	1	MALA
007M	TIFF	CCD	1	1	1	1	MALA
008M	TIFF	CCD	1	1	1	1	BUENA
009M	TIFF	CCD	1	1	1	1	REGULAR
010M	TIFF	CCD	1	1	1	1	REGULAR
011M	TIFF	CCD	1	1	0	1	MUY BUENA
012M	TIFF	CCD	1	1	1	1	MUY BUENA
013M	TIFF	CCD	1	1	1	1	MALA
014M	TIFF	CCD	1	1	1	1	BUENA
015M	TIFF	CCD	1	1	1	1	REGULAR

Se tienen en cuenta los siguientes elementos para clasificar las imágenes que harán parte de este estudio así:

- Nombre: Lleva el consecutivo de la imagen y una o dos letras que indican donde se tomó la fotografía y/o el tipo de sensor.
- Formato: Se trabajó con formatos tiff, bmp, jpg y png.

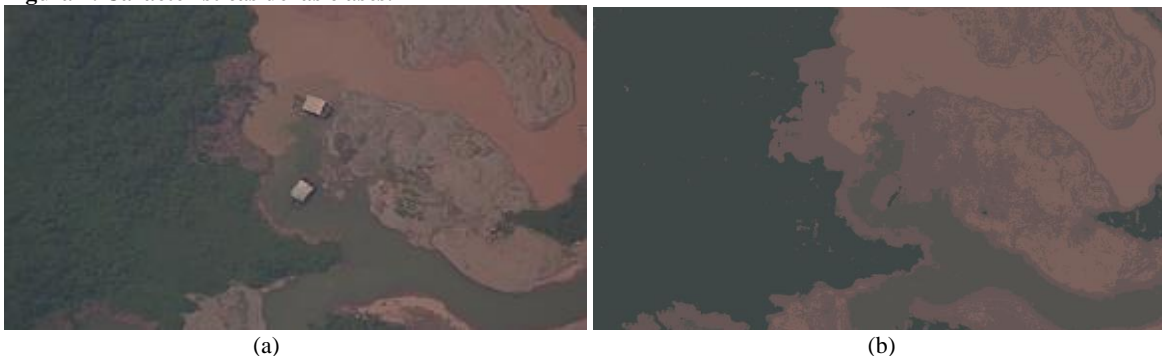
- Sensor: Se obtuvieron imágenes tomadas con sensor CCD e IR desde una plataforma aérea.
- Clases: Agua, vegetación, construcciones y otros se coloca uno (1) si la clase está presente en la imagen y cero (0) si no.
- Visión de las clases: Con ayuda de los expertos se da una calificación en cuatro escalas: Mala, regular, buena y muy buena, de acuerdo a la facilidad con que se pueden identificar las clases a simple vista.

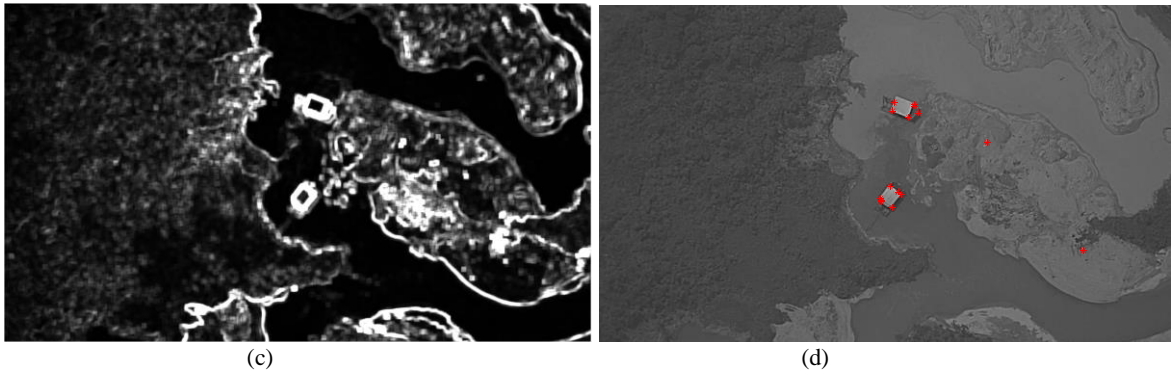
5.2 DEFINICIÓN DE CLASES

Se realizaron pruebas del comportamiento sobre los píxeles de color a través de algoritmos de clasificación no supervisada como k-means y modelo de mezcla de Gaussianas, respuesta a filtro promedio, Gaussiano y algoritmo de Canny para observar los bordes generados, detección de esquinas a través del algoritmo de Harris, elección de regiones con píxeles adyacentes diferentes con cálculo del gradiente, selección de regiones extremadamente estables (MSER) y otros algoritmos que permiten analizar el comportamiento de los píxeles y su relación con vecinos cercanos.

Con el anterior procedimiento se determina cómo las diferentes clases responden a procedimientos estadísticos que relacionan los píxeles, observándose por ejemplo que los grupos de color no reconocen las fuentes de agua, pues esta puede tener colores diferentes, sin embargo el filtro promedio, con el cual se puede analizar la varianza entre grupos de píxeles muestra que en esta misma clase, los píxeles adjuntos tienen gran probabilidad de ser iguales, reconociendo en general agua sin importar el color que tenga. De forma similar el comportamiento de los algoritmos en diferentes características muestra la importancia del contexto para determinar las clases.

Figura 4. Características de las clases.





(a) Paisaje tropical, (b) Clasificación de píxeles por color (RGB) con mezcla de Gaussianas, (c) Textura con filtro promedio (d) Detección de esquinas con algoritmo de Harris.

Finalmente se opta por analizar 4 clases para clasificación, dos que pertenecen naturalmente a los paisajes tropicales colombianos: Vegetación, Agua y dos que no pertenecen típicamente a paisajes naturales: construcciones y otros elementos en los que se pueden incluir zonas deforestadas, zonas inundadas, zonas modificadas, maquinaria, vehículos, desvío de cauce de ríos, etc.

5.3 DEFINICIÓN DE LA SEMÁNTICA DE CLASIFICACIÓN

Una vez analizadas las respuestas generadas por el procesamiento de imágenes con algoritmos no supervisados, se puede observar que las características relacionadas con el color en sus 3 planos (RGB) están definidas claramente por su media y varianza en grupos de píxeles así:

- Vegetación: Tiene diferentes tonalidades de verdes donde la media se puede ubicar, pero en un mismo grupo de píxeles divididos en cuadros de $n \times n$ dentro de esta misma clase puede cambiar varias veces su valor con respecto su media en tonalidades semejantes, es decir tiene varianza media.
- Agua: En varias imágenes se observa que puede tener diferentes tonalidades de color entre verdes, azules y cafés, pero su varianza es mínima con respecto a la media, anormalmente se pueden encontrar objetos dentro del agua, corrientes fuertes y mezcla de agua de diferentes colores que serán etiquetadas como “otros” ya que no pertenecen a un paisaje tropical normal.
- Construcciones: Tiene muchísimos colores, cafés, negros, grises y en menor cantidad azules y rojos, también tienen diferentes tamaños y no son comunes en los paisajes tropicales. En general tienen bordes uniformes (polígonos) que encierran regiones donde los píxeles cambian muy poco y

esquinas muy pronunciadas de fácil detección con un algoritmo como Harris.

- Otros: Encierra muchos comportamientos en cuanto a la media del color, su varianza es típicamente alta ya que al modificar los paisajes sanos cambian los colores abruptamente en zonas de diferentes tamaños, algunos de estos cambios son muy difíciles de identificar por el experto como lo son las zonas recientemente deforestadas o ríos que han cambiado su cauce y dejan rastros de barro, ambos casos son elementos que no pertenecen a un paisaje tropical normal.

Por lo anterior se realizó la semántica de clasificación con etiqueta de color en tres clases: vegetación, agua y otros (que incluirá las muestras de construcciones) ya que los dos primeros son elementos normales de un paisaje tropical y otros siempre serán modificaciones que deberán ser analizadas posteriormente, así una vez segmentada la imagen en estas clases se incluye la detección de otros elementos como construcciones, maquinaria, deforestación, etc., con detección a través de algoritmos que funcionan principalmente en escala de grises.

Definido lo anterior la técnica utilizada que puede hacer la transformación del espacio vectorial para reunir estos comportamientos del color en las tres clases seleccionadas utilizará los descriptores de covarianza a través del cálculo de los sigma points descritos en el numeral 2.3

5.3.1 Intervención del conocimiento experto.

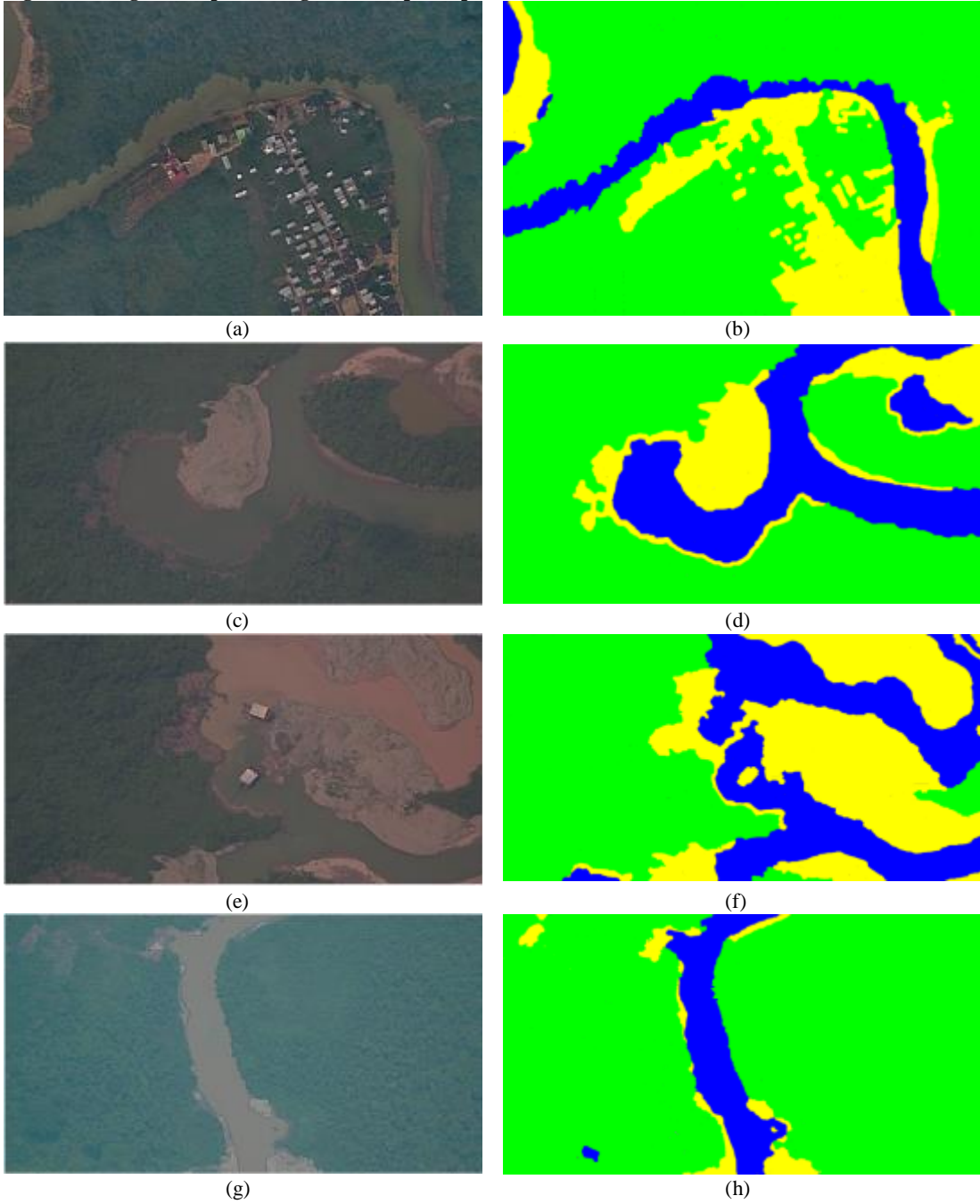
Al proponer un algoritmo de segmentación supervisado se requiere la intervención de expertos, en este caso se contó con dos personas con experiencia de 4 y 5 años respectivamente en análisis de imágenes cartográficas, quienes pudieran determinar con alto grado de certeza las zonas con cada una de las clases y que fuesen estas muestras los elementos para determinar el error en el algoritmo de clasificación y para entrenar inicialmente la máquina con la cual se pudieran obtener las imágenes completamente segmentadas.

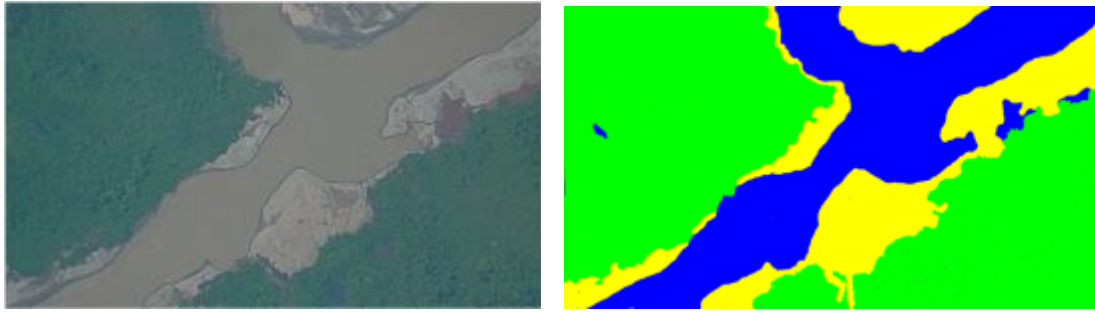
Se segmentaron de forma manual con ayuda de los analistas expertos, cinco plantillas “ground truth” de las imágenes 000MO.bmp, 011M.tif, 0012M.tif, 030M.tif y 031M.tif respectivamente como se muestra en la Figura 5. Donde el color verde corresponde a la clase vegetación, el azul al agua y el amarillo a otros.

Es imprescindible anotar que esta clasificación manual tiene una pequeña probabilidad de tener algún error sistemático en los bordes e identificación de algunas sombras que pueden ser confundidas, sin embargo, se tomarán precauciones al tomar las muestras de entrenamiento para que esas zonas no

pertenezcan a los datos de prueba, así, estas imágenes serán tomadas como la plantilla 100% correcta en la clasificación de píxeles.

Figura 5. Imágenes de prueba segmentadas por expertos.





(i)

(j)

(a) 000MO.bmp, (b) 000MOGT.bmp (c) 011M.tif, (d) 011MGT.bmp, (e) 0012M.tif, (f) 0012MGT.bmp, (g) 030M.tif, (h) 030MGT.bmp, (i) 031M.tif, (j) 031M.bmp

5.4 ALGORITMO DE CLASIFICACIÓN

Para el desarrollo y explicación del algoritmo solamente se toma en cuenta la imagen 012M.tif, la cual tiene las clases requeridas, en el Anexo 1 se encuentra la evaluación del algoritmo con las 5 imágenes ground truth segmentadas por el experto a las cuales se les hicieron las mismas pruebas en condiciones similares.

El algoritmo se basa en encontrar la semántica de grupos de píxeles correspondientes a cada clase definida para clasificar otros grupos dentro de la misma imagen o en imágenes similares, de esta forma es posible segmentar una imagen completa con pocas muestras de sí misma o alimentar una base de datos para el aprendizaje de máquina que permita mejorar la calidad en el reconocimiento de clases y reducir el error de clasificación.

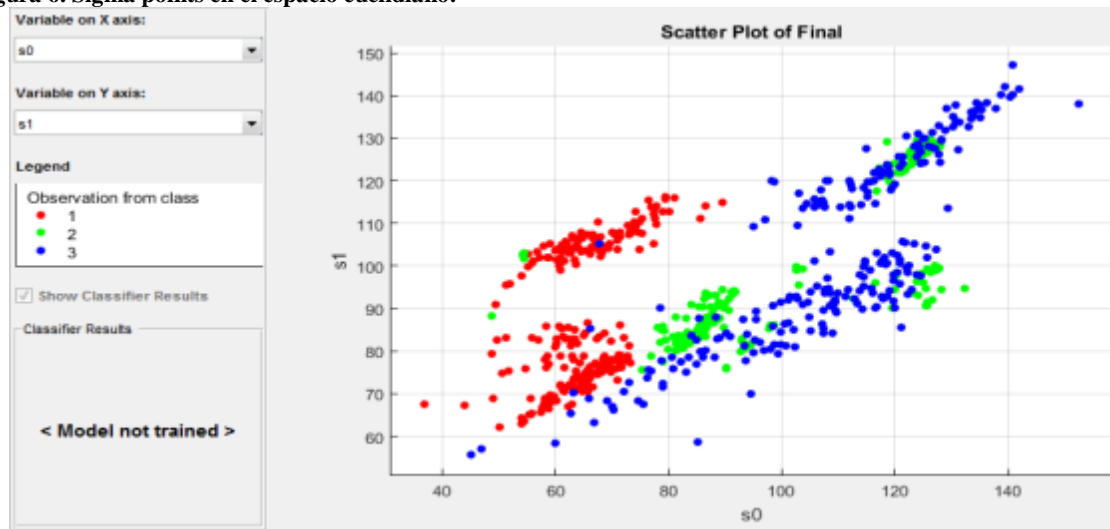
5.4.1 Selección de parámetros de prueba

A partir de la plantilla ground truth clasificada por los expertos se obtuvieron muestras las imágenes 000MO.bmp, 011M.tif, 012M.tif, 030M.tif y 031M.tif, en grupos de 10, 25 y 50 muestras de cada clase por imagen, teniendo cuidado de no muestrear áreas con sombras o bordes entre las clases vegetación, agua y otros. La selección de muestras es configurable en cuadros de $n \times n$, seleccionando cuadros de 10×10 debido a la forma y tamaño de los objetos, evitando las sombras y bordes entre clases, se pueden seleccionar cuadros más grandes reduciendo el tiempo de procesamiento, pero aumentando el error de clasificación. Todas las muestras son recopiladas en una tabla, obteniendo en total 750 muestras, 250 por cada clase.

La presentación de los datos para el clasificador se realizó a través de datos de color (RGB) procesados para encontrar los “*sigma points*” de acuerdo a la metodología detallada en el numeral 2.4.1 aplicada a la población inicial de 750

muestras. Los datos etiquetados con 21 características son mostrados en la figura 6.

Figura 6. Sigma points en el espacio euclidiano.



En rojo se observan los puntos de la clase vegetación, en verde agua y en azul otros.

El método de entrenamiento de máquina se escogió teniendo en cuenta su aplicación a un método diferente al de árbol de decisión el cual ya fue probado de acuerdo a la bibliografía consultada [9]. También fue evaluada la simplicidad, posibilidad de exportar los modelos a un lenguaje de programación orientada a objetos para su implementación, menor coste computacional, menor error en entrenamiento y menor error de clasificación.

Se escogió SVM ya que cumple la mayor parte de estas características [3] pero, al no aplicarse a un clasificador binario se requiere evaluar los tipos de kernel para mejores resultados, este estudio se hace con el toolbox de Matlab mostrado en la figura 7 haciendo el cálculo del error de entrenamiento y luego evaluando el de clasificación.

Figura 7. Entrenamiento SVM toolbox Matlab con 750 muestras, diferente kernel

SVM	
Linear SVM	97.6%
SVM	
Quadratic SVM	98.8%
SVM	
Cubic SVM	100%
SVM	
Fine Gaussian SVM	100%

Observando que las muestras son cuasi-perfectamente diferenciables en todos los casos mostrando los mejores resultados con kernel cúbico y gaussiano fino, se procede a encontrar el error de clasificación para cada uno aplicando los modelos en la totalidad de la imagen 012M.tif y comparándola con la plantilla.

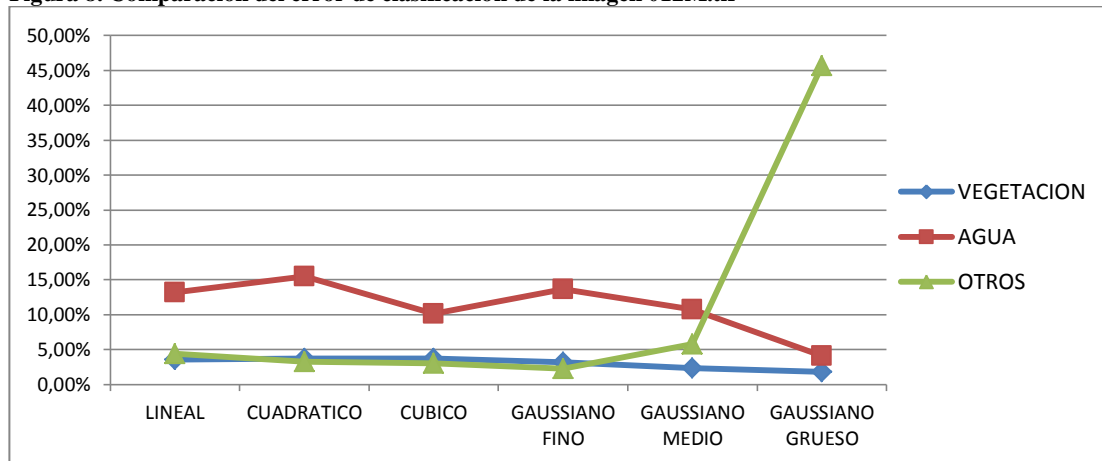
Tabla 4. Error de clasificación con 750 muestras, imagen 012M.tif

KERNEL	LINEAL	CUADRATICO	CUBICO	GAUSSIANO FINO	GAUSSIANO MEDIO	GAUSSIANO GRUESO
Vegetación	3,58%	3,69%	3,75%	3,20%	2,32%	1,81%
Agua	13,20%	15,53%	10,16%	13,63%	10,74%	4,13%
Otros	4,43%	3,29%	3,02%	2,24%	5,82%	45,68%

En rojo se muestran los peores casos de clasificación y en gris los mejores.

La figura 8 muestra el comportamiento del error de clasificación de cada clase, se debe tener en cuenta que la prioridad para el usuario final es encontrar la clase otros en un paisaje tropical, por lo tanto se opta por continuar las pruebas con SVM Fine Gaussian.

Figura 8. Comparación del error de clasificación de la imagen 012M.tif



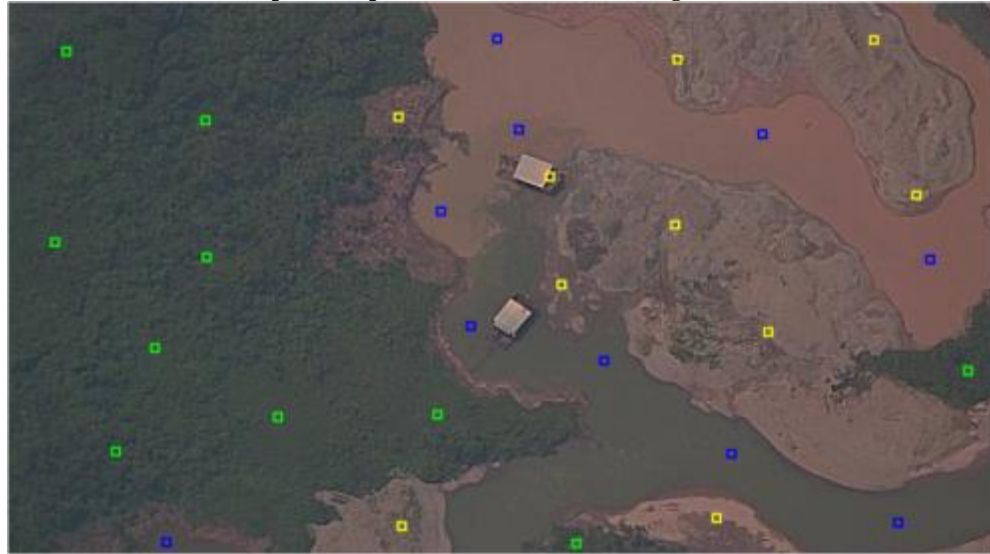
5.4.2 Selección de los datos de entrenamiento

Para el análisis se escogen secciones cuadradas variables, se debe tener en cuenta que entre más grande el cuadro, aumenta la velocidad de procesamiento pero clasifica con menor precisión. Por esta razón se tomaron cuadros de 10x10 seleccionados aleatoriamente dentro de las zonas segmentadas por la imagen ground-truth.

Primero se seleccionan las muestras de cada una de las clases para entrenamiento, inicialmente con 10, se agregan 15 para tener 25 y finalmente se suman 25 más para tener 50 cuadros por imagen, de esa forma se puede

observar el comportamiento del error frente a la cantidad de muestras de entrenamiento. El proceso se muestra en la figura 9.

Figura 9. Selección de 10 muestras por clase para entrenamiento de la imagen 012M.tif



5.4.3 Error de muestras por cuadros de entrenamiento

Los datos tanto de entrenamiento como los de los resultados de clasificación se guardan en una tabla para su uso posterior, las muestras se tomaron aleatoriamente pero en forma supervisada para seleccionarlas secuencialmente y encontrar por simple comparación los errores de los datos por cuadros de prueba, también se revisó la matriz de confusión con el fin de analizar los resultados frente al error por cada clase.

De la imagen 011M.tif se toman 50 muestras por cada clase para un total de 150 que son entrenadas con SVM kernel *“fine gaussian”* el modelo se aplica a 30 cuadros aleatorios (10 por cada clase) de la misma imagen (011M.tif) y de otra imagen (012M.tif) de esta forma puede probarse la base del método calculando el error cuadro por cuadro comparándolos según su posición en las imágenes groundtruth, los resultados se observan en la Tabla 5.

Tabla 5. Error de muestras (cuadros) para entrenamiento

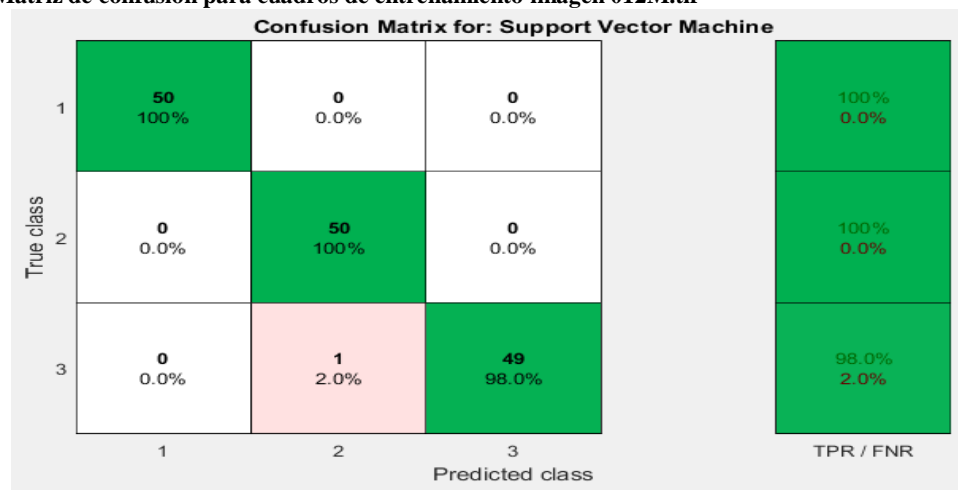
No. Muestra	012M Clase1	011M Clase1	ECM Clase1	012M Clase2	011M Clase2	ECM Clase2	012M Clase3	011M Clase3	ECM Clase3
1	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0
2	Veg	Agua	1	Agua	Agua	0	Otro	Otro	0
3	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0
4	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0
5	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0

6	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0
7	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0
8	Veg	Agua	1	Agua	Agua	0	Otro	Otro	0
9	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0
10	Veg	Veg	0	Agua	Agua	0	Otro	Otro	0

Error cuadrático medio para muestras de la misma imagen de donde provienen las muestras de entrenamiento es 0% y el de otra imagen se calcula en el 6,66%

La matriz de confusión de la tabla 6 da una idea de donde se producen los errores entre clases, producto del entrenamiento.

Tabla 6. Matriz de confusión para cuadros de entrenamiento imagen 012M.tif



La matriz muestra que la probabilidad de error se presenta entre las clases 3(otros) y 2(Agua), según el conocimiento de un experto, para un analista de imágenes es complicado diferenciar entre una playa producto de la sequía de un río y la misma fuente de agua puesto que la profundidad en las orillas es baja. Sin embargo, reconocer este patrón resulta muy importante cuando se quiere determinar el impacto debido al desbordamiento de los ríos y contaminación de los mismos.

5.4.4 Pruebas en imagen de entrenamiento

El desarrollo y puesta a punto del algoritmo de segmentación se hizo con la imagen 012M.tif, el método escogido para hacer el análisis de error consistió en obtener resultados de la imagen segmentada completamente con 10, 25 y 50 muestras para entrenamiento con la certeza del experto de pertenecer a cada clase. Luego se divide la imagen completa en una malla de cuadros de 10 x 10 como se observa en la figura 10 para clasificar la totalidad de los cuadros según el

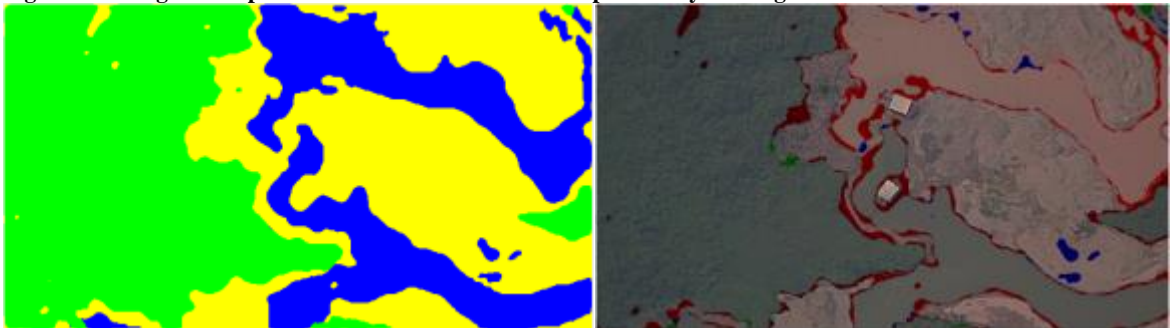
modelo de aprendizaje SVM de kernel *“fine gaussian”*. Se emplea esta técnica para reducir el tiempo de procesamiento sin aumentar demasiado el error de clasificación.

Figura 10. Selección de cuadros de 10 x 10 para clasificación de la imagen 012M.tif

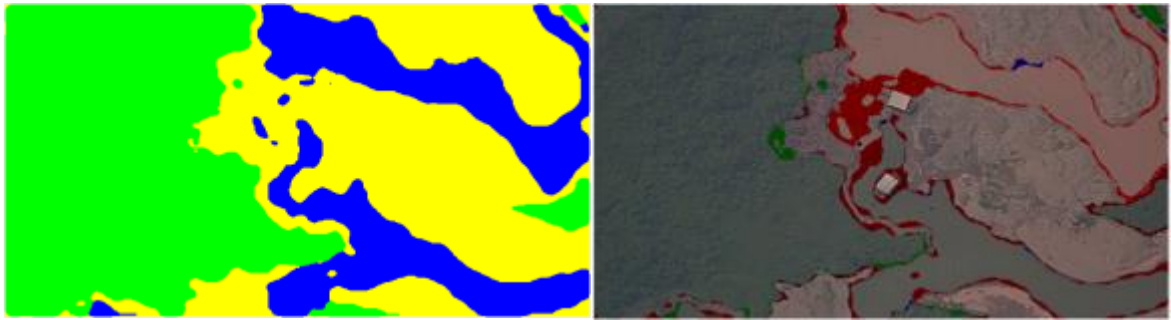


Cada cuadro se pinta con alguno de los colores que representan las clases: verde para vegetación, azul para agua y amarillo para otros, con el fin de recuperar los bordes entre clase se aplica un filtro de moda del doble del tamaño del cuadro clasificado, finalmente la clasificación a partir de 10, 25 y 50 muestras respectivamente se compara con el modelo obtenido de la imagen *“ground truth”* obteniendo los resultados que se muestran en la figura 11.

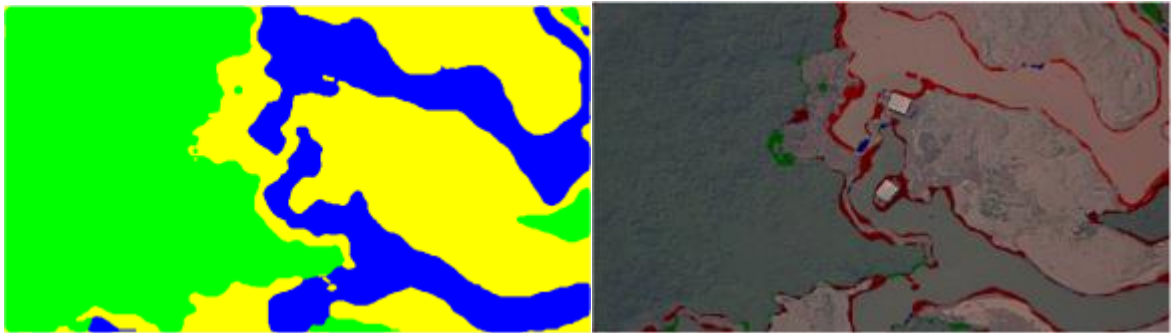
Figura 11. Imágenes de prueba clasificadas con muestras por clase y su imagen de error.



(a)



(b)



(c)

(a) 10 muestras comparado con la ground truth.(b) 25 muestras comparado con la ground truth, (c) 50 muestras comparado con la ground truth. Los pixeles mal clasificados se muestran en verde azul y rojo en la imagen original.

La matriz de confusión para el error de entrenamiento en 50 muestras muestra cómo se distribuye el error entre las clases.

Tabla 7. Matriz de confusión para imagen 012M.tif clasificada con 50 muestras de ella misma

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	97,1%	0,0%	2,9%	42%	97%	3%
2	0,6%	85,1%	14,3%	25%	85%	15%
4	1,8%	0,5%	97,7%	33%	98%	2%
Total				100%		

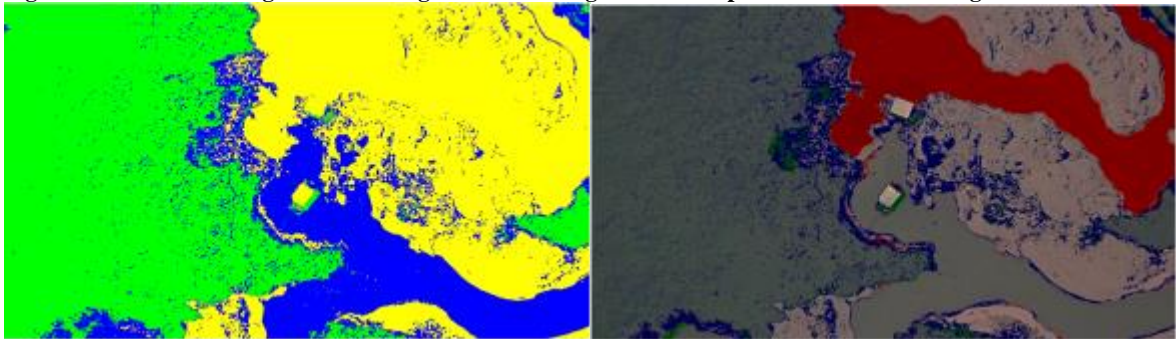
En esta imagen la principal confusión se presenta entre agua y otros, esto sucede porque hay zonas modificadas por el hombre producto de la minería ilegal donde se desvían cauces de ríos y se mueven volúmenes de tierra que al final terminan siendo barro, el cual puede ser confundido con agua del mismo río. Esta imagen también muestra la combinación de agua contaminada de procesos químicos para extracción de metales con agua de un río, esto claramente no es normal y las zonas donde ocurre esto deben ser identificadas para hacer seguimiento del impacto económico, ambiental y social.

5.4.5 Pruebas de imagen segmentada con algoritmo por píxel

Ahora se realiza la misma prueba de error con la imagen 012M.tif, pero esta vez segmentada con el algoritmo no supervisado de mezcla de gaussianas mostrado en la figura 12, en este caso se evalúa cada píxel y se clasifica en un grupo de $N = 3$ distribuciones gaussianas que representan cada clase con una media específica, los píxeles se estiman con la mayor probabilidad de pertenecer a alguna clase cercana a su media de color en sus componentes RGB.

Se escoge este algoritmo de comparación debido a la similitud en encontrar información estadística en los píxeles, así se puede emitir un juicio con mayor objetividad sobre la efectividad del algoritmo de segmentación semántica.

Figura 12. Resultado imagen 012M.tif segmentada con algoritmo no supervisado de mezcla de gaussianas



La matriz de confusión muestra alta probabilidad de error entre clase agua y otros ya que el algoritmo reúne solamente colores y no relación entre colores cercanos, en este caso una solución es agregar una distribución más ($N=4$) para separar el color del barro con otro color de agua.

Tabla 8. Matriz de confusión para imagen 012M.tif clasificada con GMM

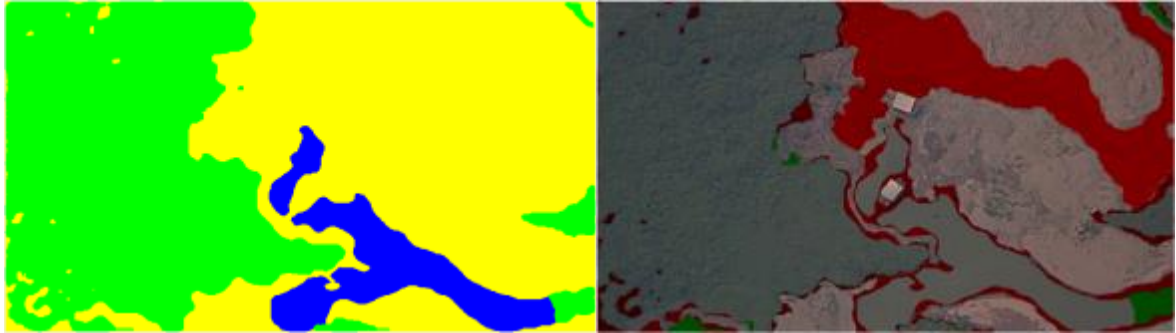
TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	94,1%	5,6%	0,3%	42%	94%	6%
2	0,6%	46,7%	52,7%	25%	47%	53%
4	2,5%	21,7%	75,8%	33%	76%	24%
TotalxClase				100%		

5.4.6 Pruebas y cálculo de error en imagen entrenada con datos de otra

Uno de los requerimiento de usuario final es poder enriquecer la base de datos para segmentar nuevas imágenes con datos existentes, para comprobar si el algoritmo mejora la calidad en la segmentación usando información recopilada

anteriormente, se prueba con 50 muestras de entrenamiento obtenidas de la imagen 011M.tif, luego se aplica el modelo obtenido a la imagen 012M.tif y se procede a segmentar la imagen completa empleando el método ya descrito, el resultado se observa en la figura 13, al final las muestras de ambas imágenes enriquecerán la población total para evaluación de nuevas entradas.

Figura 13. Resultado imagen 012M.tif segmentada con 50 muestras de imagen 011M.tif



El resultado es muy favorable teniendo en cuenta que la mayor parte del error se encuentra en el agua como se muestra en la tabla 6, la cual, en la imagen de entrenamiento 011M.tif tiene dos tonalidades de color pero ninguna coincide con una de las tonalidades de color de agua de la imagen 012M.tif. La metodología con prioridad a la clase otros, permite entonces enfocar un análisis posterior en estas zonas para mejorar la segmentación resultante y tomar nuevas muestras de las clases definidas.

Tabla 9. Matriz de confusión para imagen 012M.tif clasificada con 50 muestras de la imagen 011M.tif

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	93,6%	0,0%	6,3%	42%	94%	6%
2	3,6%	32,3%	64,1%	25%	32%	68%
4	0,8%	0,0%	99,2%	33%	99%	1%
TotalxClase				100%		

5.4.7 Pruebas y cálculo de error en imagen clasificada con GMM de otra

Finalmente para comparar el rendimiento del algoritmo semántico supervisado, se realiza la clasificación de los píxeles de la imagen 012m.tif, con el modelo gaussiano generado de la imagen 011M.tif, el resultado se muestra en la figura 14.

La matriz de confusión muestra un resultado muy diferente en cuanto a los píxeles clasificados pues un alto porcentaje de píxeles que pertenecen a las clases agua y otros fueron confundidas con vegetación, como se muestra en la Tabla 7.

Figura 14. Resultado imagen 012M.tif segmentada con GMM de imagen 011M.tif

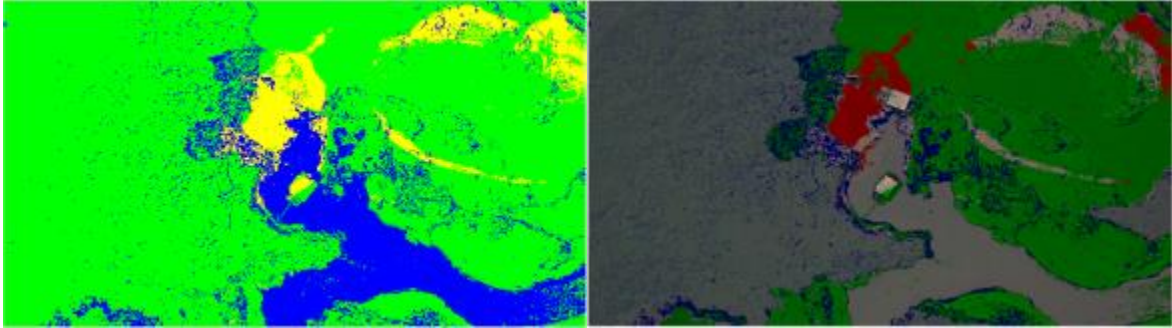


Tabla 10. Matriz de confusión para imagen 012M.tif clasificada GMM de la imagen 011M.tif

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	96,5%	3,4%	0,0%	42%	97%	3%
2	45,7%	43,5%	10,9%	25%	43%	57%
4	72,1%	16,0%	11,9%	33%	12%	88%
TotalxClase				100%		

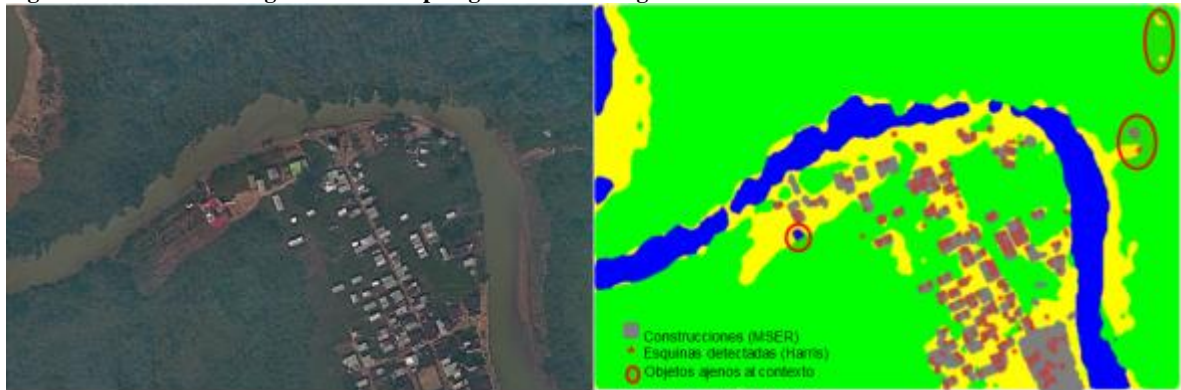
5.5 PRODUCTO FINAL DEL ALGORITMO DE SEGMENTACIÓN SEMÁNTICA

En la práctica una misión aérea de toma de fotografías puede contener cientos de fotografías de áreas extensas de terreno que deben ser evaluadas para analizar, identificar y tomar decisiones sobre los objetos encontrados en ellos. El significado de estos objetos puede ser variado, desde el impacto ambiental que produce el hombre en la selva tropical, hasta la prevención de desastres en zonas pobladas. Actualmente se cuenta con un software para realizar mosaicos ortogonales a partir de una gran cantidad de fotos, estas imágenes resultantes pueden ser procesados con el algoritmo semántico para encontrar las clases de interés y a partir de ellas identificar elementos que a simple vista son difícilmente detectables, o para hacer mediciones cartográficas de áreas de bosque, crecimiento de ríos o desviación de cauces, lo anterior teniendo en cuenta que se dispone de georreferenciación en las imágenes.

Con la segmentación se observan zonas claramente diferenciadas en 3 de las cuatro clases definidas inicialmente, luego se buscan regiones extremadamente estables (MSER regions) las cuales son la característica de las edificaciones, en este caso se espera encontrar esta cuarta clase en una zona de color amarillo y clasificada en otros, encontrarlas fuera de estas zonas requiere atención especial pues en el contexto de los paisajes tropicales esta clase no debe estar dentro de regiones clasificadas como vegetación o agua. De la misma forma, a través del algoritmo de Harris pueden buscarse esquinas o extremos filosos los cuales

generalmente pertenecen a objetos fabricados por el hombre como construcciones, vehículos o maquinaria que tampoco deben hacer parte del contexto del paisaje tropical.

Figura 15. Resultado imagen 000OM.bmp segmentada con algoritmo semántico

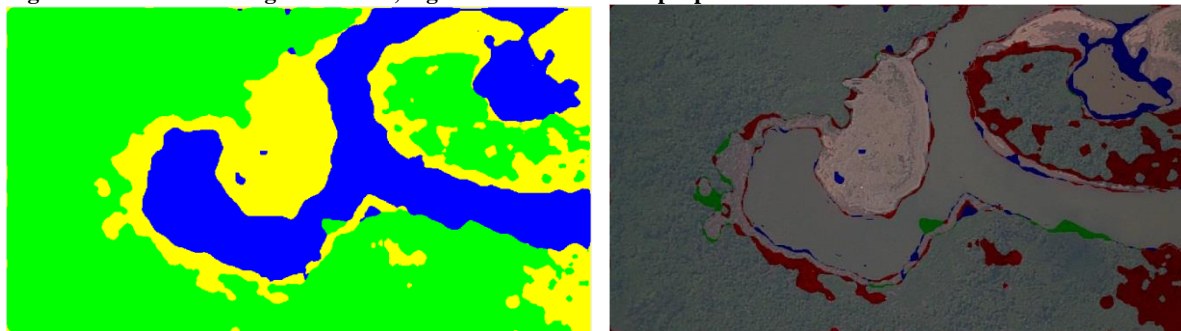


Es analizado el resultado con un experto, se encuentran 5 objetos de especial atención que son ajenos al contexto de las clases segmentadas.

5.5.1 Pruebas y cálculo de error del método propuesto caso I

Se realiza una prueba del método, la cual consiste en utilizar la base de datos existente y agregar imágenes de cualquier tipo, formato o toma que permita conocer la eficiencia frente a nuevos datos el proceso toma como base de datos inicial las 50 muestras de la imagen 000MO.bmp, se toman 10 muestras de la imágenes siguientes, es decir, desde la 001M.tif hasta la 010M.tif, se aplica el resultado del entrenamiento de 450 muestras de 10 x 10 con la imagen 011M.tif provenientes de las imágenes anteriores con SVM fine Gaussian, los resultados obtenidos y a prueba de error se muestran en la figura 16.

Figura 16. Resultado imagen 011M.tif, segmentada con método propuesto



Puede observarse que con la máquina de aprendizaje pueden obtenerse resultados de segmentación con muy bajo error en la clasificación.

Tabla 11. Matriz de confusión imagen 011M.tif segmentada con 450 muestras de otras imágenes

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	87,3%	0,2%	12,5%	61%	87%	13%
2	1,6%	95,2%	3,2%	22%	95%	5%
4	2,9%	10,9%	86,3%	17%	86%	14%
TotalxClase				100%		

En este caso la clasificación correcta de la imagen en cada clase pasa del 85%, que según el experto es suficiente para emitir un concepto de valor sobre el análisis de terreno, sin embargo, un caso que puede presentarse en el método consiste en que aunque se tengan muchas muestras, no sean suficientes para clasificar correctamente grupos específicos de píxeles en la clase correcta. En este caso el algoritmo permite mejorar los resultados tomando una pequeña cantidad de muestras de la imagen por parte del experto que permita mejorar los resultados, reduciendo el tiempo de análisis.

Aunque estos resultados son relevantes para el análisis de imágenes es posible mejorar los resultados haciendo uso de la inteligencia artificial del algoritmo, adicionando algunas muestras de la imagen original en zonas donde el analista tiene certeza de las clases, en el ejemplo de la figura 17 se adicionan 10 muestras de las ya disponibles de las pruebas anteriores.

Figura 17. Resultado imagen 011M.tif, mejorada con muestras de sí misma

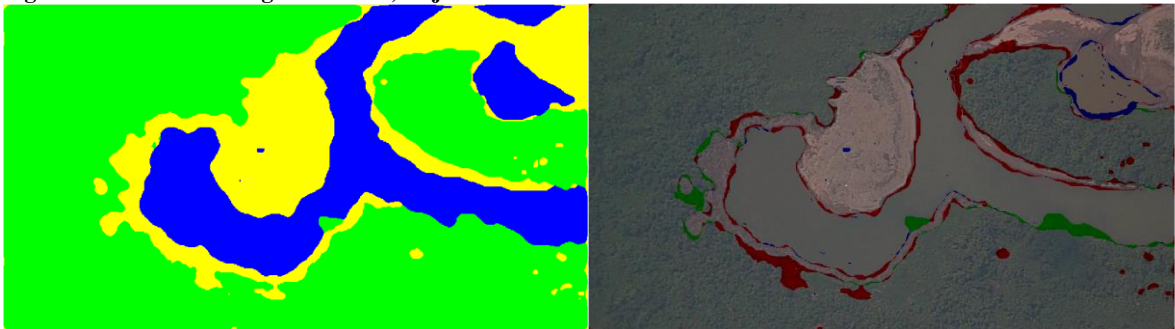


Tabla 12. Matriz de confusión imagen 011M.tif mejorada con muestras de sí misma

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	96,1%	0,0%	3,9%	61%	96%	4%
2	2,8%	91,2%	6,0%	22%	91%	9%
4	3,8%	3,5%	92,7%	17%	93%	7%
TotalxClase				100%		

5.5.2 Pruebas y cálculo de error del método propuesto caso II

Finalmente se hace una prueba de segmentación que incluye una mayor cantidad de muestras provenientes de imágenes de la base de datos, se obtienen al menos 10 muestras de cada clase con ayuda de un experto de las imágenes 000MO.bmp, 001M.tif, 002M.tif, 003M.tif, 004M.tif, 005M.tif, 006M.tif, 007M.tif, 008M.tif, 009M.tif, 010M.tif, 011M.tif, 014M.tif, 015M.tif, 016M.tif, 017M.tif, 018M.tif, 019M.tif, 020M.tif, 021M.tif, 022M.tif, 023M.tif, 024M.tif, 025M.tif, 026M.tif, 027M.tif, 028M.tif, 029M.tif, 030M.tif, 031M.tif, 032M.tif, 033M.tif, 034M.tif, 035M.tif, 036M.tif, 039M.tif, 040M.tif, 041M.tif y 042M.tif, obteniendo un total de 960 muestras para entrenamiento que se aplica en la imagen 012M.tif de la cual no se obtuvo hasta ahora ninguna muestra. El resultado se muestra en la figura 18 y tabla 9.

Figura 18. Resultado imagen 012M.tif, segmentada 960 muestras de 39 imágenes

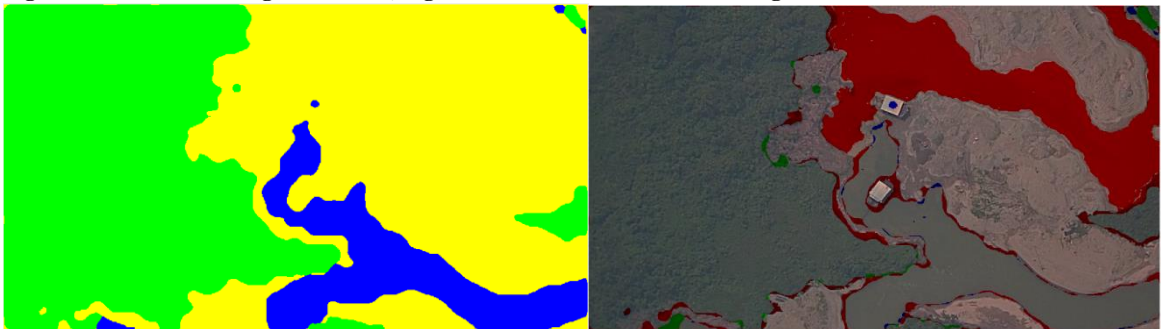


Tabla 13. Matriz de confusión imagen 012M.tif segmentada 960 muestras de 39 imágenes

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	96,3%	0,0%	3,7%	42%	96%	4%
2	1,0%	41,7%	57,3%	25%	42%	58%
4	1,1%	0,6%	98,3%	33%	98%	2%
TotalxClase				100%		

Figura 19. Resultado imagen 012M.tif, mejorada con muestras de sí misma

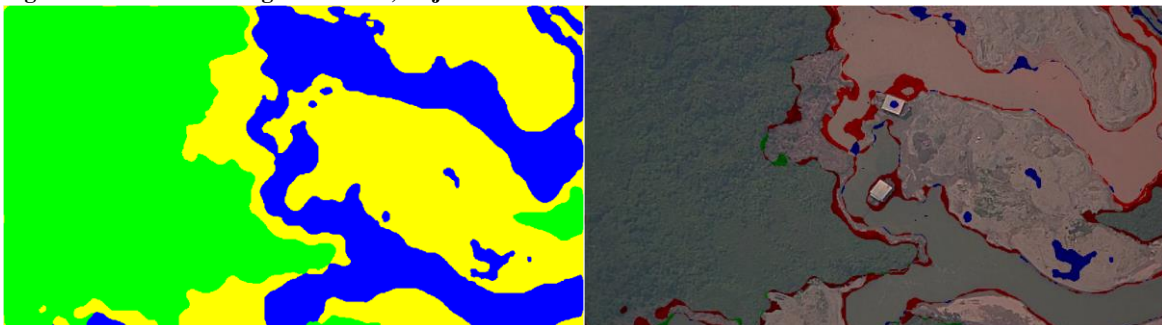


Tabla 14. Matriz de confusión imagen 012M.tif, mejorada con muestras de sí misma

TRUE CLASS	CLASE PREDECIDA			Total x Clase	Píxeles Correctos	Píxeles erróneos
	1	2	4			
1	95,6%	0,0%	4,4%	42%	96%	4%
2	0,2%	87,8%	12,0%	25%	88%	12%
4	0,7%	3,7%	95,6%	33%	96%	4%
TotalxClase				100%		

6. ANÁLISIS DE RESULTADOS

Se analizan los resultados del algoritmo de segmentación semántica comparados con el algoritmo de segmentación por pixel de mezcla de gaussianas (GMM) para imágenes de paisajes tropicales en sus componentes de color RGB obtenidas con cámara CCD para imágenes tropicales. No se realizó análisis semántico en imágenes de sensores infrarrojos debido a que se obtuvieron muy pocas fotografías IR y la mayoría no eran de paisajes tropicales. Por esta razón no fue posible encontrar patrones de comparación para aplicar el algoritmo. De la misma forma la aplicación de sigma points resulta irrelevante en imágenes obtenidas con este sensor infrarrojo para análisis de color puesto que las imágenes son en escala de grises.

6.1 ANÁLISIS DE ERROR PARA EL ALGORITMO DE SEGMENTACIÓN

La tabla 8 muestra la recopilación de resultados para las pruebas realizadas con las 5 imágenes seleccionadas y comparadas con la plantilla ground truth clasificada por un experto. La relación de píxeles clasificados correctamente con cada método se muestra con sombra gris y el peor caso con sombra roja, también se relaciona el porcentaje de píxeles pertenecientes a cada clase que hay en la imagen evaluada. Las imágenes segmentadas en cada caso se pueden observar en el Anexo 1.

Tabla 15. Píxeles correctamente clasificados en 5 imágenes evaluadas con información de sí mismas

INFORMACION DE LAS IMÁGENES			ANÁLISIS DE ERROR			
CLASE	IMAGEN	Píxeles x Clase	10	25	50	GMM
VEGETACIÓN	000MO.bmp	70%	55,53%	76,00%	88,25%	95,70%
	011M.tif	61%	91,73%	93,45%	95,07%	91,70%
	012M.tif	42%	95,11%	97,58%	97,10%	94,10%
	030M.tif	88%	62,01%	95,88%	98,60%	75,50%
	031M.tif	54%	88,57%	92,67%	93,99%	82,30%
AGUA	000MO.bmp	11%	87,96%	73,86%	73,99%	95,50%
	011M.tif	22%	65,76%	79,61%	84,10%	95,60%
	012M.tif	25%	86,44%	81,28%	85,10%	46,70%
	030M.tif	9%	84,34%	84,99%	85,55%	98,50%
	031M.tif	26%	70,06%	91,28%	92,20%	99,80%
OTROS	000MO.bmp	19%	95,50%	96,93%	97,24%	55,60%
	011M.tif	17%	97,84%	97,03%	95,77%	61,50%
	012M.tif	33%	97,34%	97,32%	97,73%	75,80%

	030M.tif	4%	96,19%	92,63%	86,35%	3,30%
	031M.tif	20%	97,92%	97,72%	97,02%	0,40%

En general se comprueba que con un porcentaje muy pequeño de muestras para entrenamiento en cada imagen para segmentación semántica se pueden obtener resultados correctos, por encima del 55%, mediante una técnica simple de entender para un usuario final, quien solo tendrá que hacer clic en donde tenga la certeza que es muestra de la clase que desea clasificar. El algoritmo permite además tomar mayor cantidad de muestras para obtener precisión en la clasificación por encima del 85%, segmentando eficientemente las imágenes por completo y guardando las muestras para análisis futuros.

En la técnica de mezcla de gaussianas se pueden obtener muy buenos resultados sobre todo para la clase Agua siempre y cuando las tonalidades sean similares, esto se puede solucionar agregando más distribuciones al análisis, sin embargo este proceso no es simple ni intuitivo para el usuario final quien tendría que seleccionar algún offset en la media o varianza para mejorar la clasificación. Por otro lado, la clase “otros” presenta errores muy altos, hasta del 99% debido a que esta clase requiere evaluar con más características la relación entre píxeles cercanos.

Puede observarse también que con 50 muestras de cada clase para un total de 150 muestras por imagen que corresponde 0,16% del 100% de cuadros de 10x10 disponibles para una imagen de 1290 x 720 píxeles, son suficientes para obtener una segmentación muy precisa, por encima del 84% de los píxeles clasificados correctamente y llegando hasta el 99% de efectividad.

6.2 ANÁLISIS DE ERROR PARA EL MÉTODO PROPUESTO

Como se discutió anteriormente el objetivo para el usuario final es poder hacer un análisis previo de imágenes nuevas para hacer estudios más profundos, la técnica requiere poner a prueba imágenes nuevas con información previa de tal forma que sobre las zonas de duda se haga un análisis más profundo. Se comparan los resultados para el caso de la segmentación semántica versus la segmentación por píxeles GMM, en este caso hay que destacar que el análisis en ambos casos es supervisado pues se toman decisiones a partir de la información previa. La comparación de los resultados se puede observar en la tabla 9. También se observan los resultados finales de una imagen segmentada con algoritmo semántico y todos los datos disponibles, se observa la eficiencia del algoritmo a medida que la cantidad de muestras aumenta.

Tabla 16. Píxeles correctamente clasificados en 5 imágenes con información de otra imagen similar

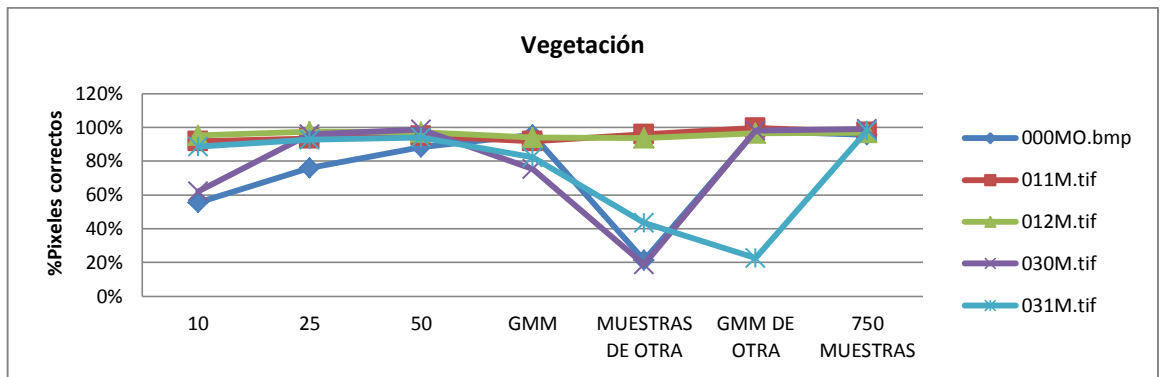
INFORMACION DE LAS IMÁGENES			ANÁLISIS DE ERROR DE CLASIFICACIÓN CON NUEVAS IMÁGENES		PÍXELES BIEN CLASIFICADOS CON TODOS LOS DATOS
CLASE	IMAGEN	Píxeles x Clase	50 MUESTRAS DE OTRA	GMM DE OTRA	750 MUESTRAS
VEGETACION	000MO.bmp	70%	21,31%	97,72%	95,45%
	011M.tif	61%	95,90%	99,79%	97,10%
	012M.tif	42%	93,60%	96,53%	96,80%
	030M.tif	88%	18,70%	98,14%	98,99%
	031M.tif	54%	43,50%	22,79%	98,19%
AGUA	000MO.bmp	11%	24,95%	7,91%	84,13%
	011M.tif	22%	82,40%	0,00%	92,56%
	012M.tif	25%	32,30%	43,45%	86,37%
	030M.tif	9%	38,70%	34,52%	88,52%
	031M.tif	26%	0,40%	99,94%	93,67%
OTROS	000MO.bmp	19%	94,77%	0,00%	89,94%
	011M.tif	17%	95,10%	17,80%	91,73%
	012M.tif	33%	99,20%	11,87%	97,76%
	030M.tif	4%	99,90%	90,92%	84,76%
	031M.tif	20%	96,90%	1,36%	94,32%

En general se observan mejores resultados en segmentación semántica principalmente para la clase otros, para el usuario final es más valioso tener precisión en esta clase, ya que el objetivo de búsqueda es encontrar anomalías en los paisajes tropicales que normalmente solo deben estar compuestos de las clases vegetación y agua. Por esta razón el algoritmo semántico se adapta mejor a estas condiciones, es más simple de usar y tiene mayor efectividad clasificando los píxeles desconocidos.

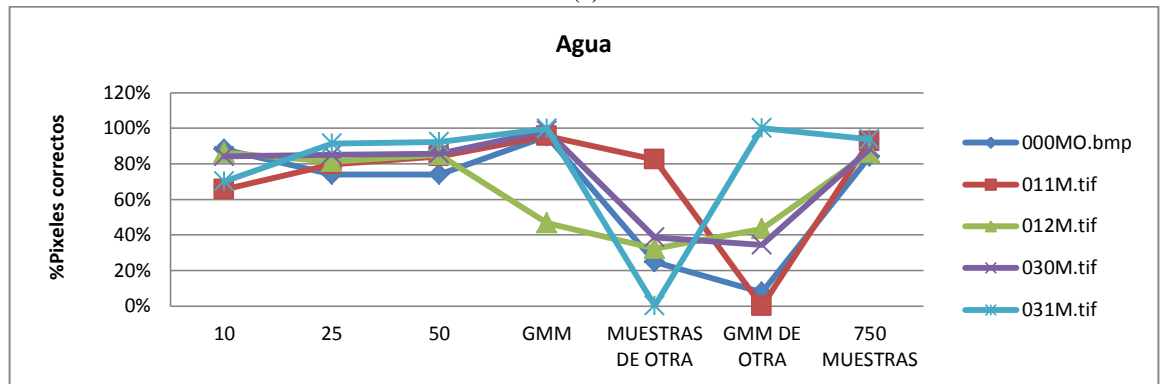
6.2.1 Análisis de resultados por clase

En cuanto a la clasificación de píxeles por clase se observa que en la segmentación semántica el porcentaje clasificado correctamente en todas las clases aumenta en cuanto más muestras se tengan sin influir enormemente una clase con respecto a las otras. Mientras que en la segmentación con mezcla de gaussianas el error en la clase agua es mínimo pero en detrimento de las demás principalmente en otros como se puede observar en la figura 16.

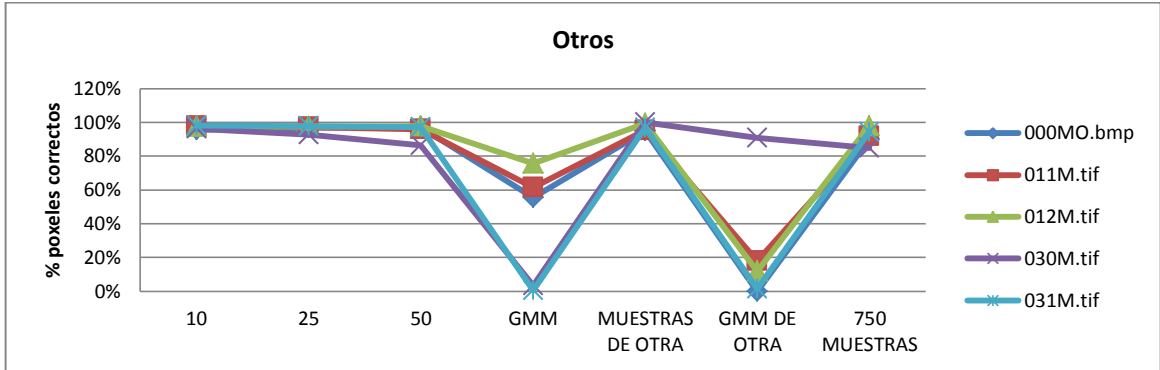
Figura 20. Píxeles clasificados correctamente vs algoritmo de segmentación aplicado



(a)



(b)

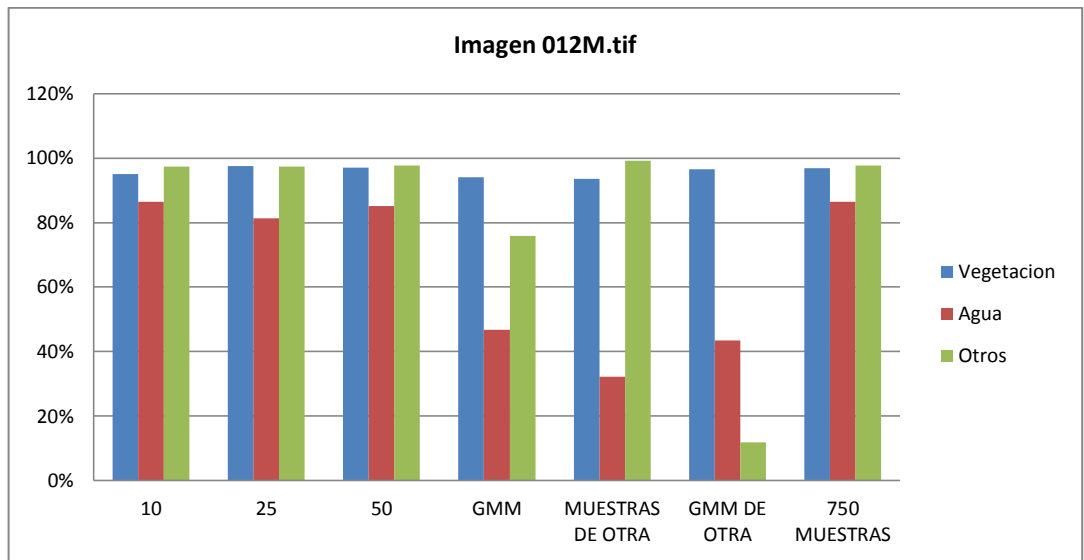


(c)

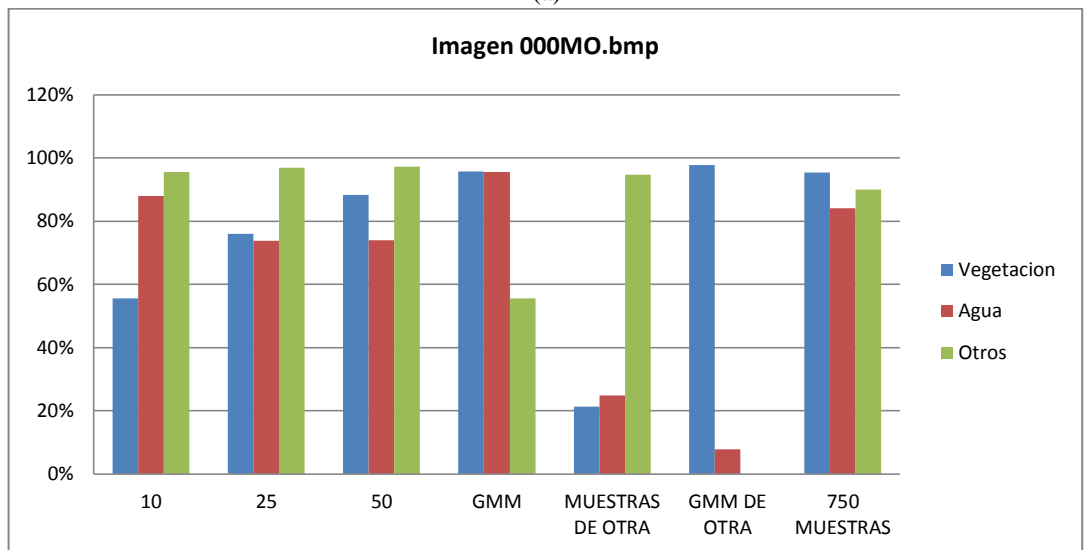
(a) Vegetación, (b) Agua, (c) Otros

Siempre que las tres clases segmentadas estén por encima del 80% de clasificación correcta, la segmentación permite tomar decisiones de importancia sobre la cartografía estudiada. Así analizando dos de las imágenes estudiadas hasta ahora en la figura 17 se observa que el algoritmo de segmentación semántica es un método muy eficiente y ampliamente superior en sus resultados que la clasificación por píxeles, para segmentar imágenes de paisajes tropicales.

Figura 21. Píxeles clasificados correctamente vs algoritmo de segmentación



(a)



(b)

(a) Aplicado a vegetación, (b) Aplicado a agua, (c) Aplicado a otros

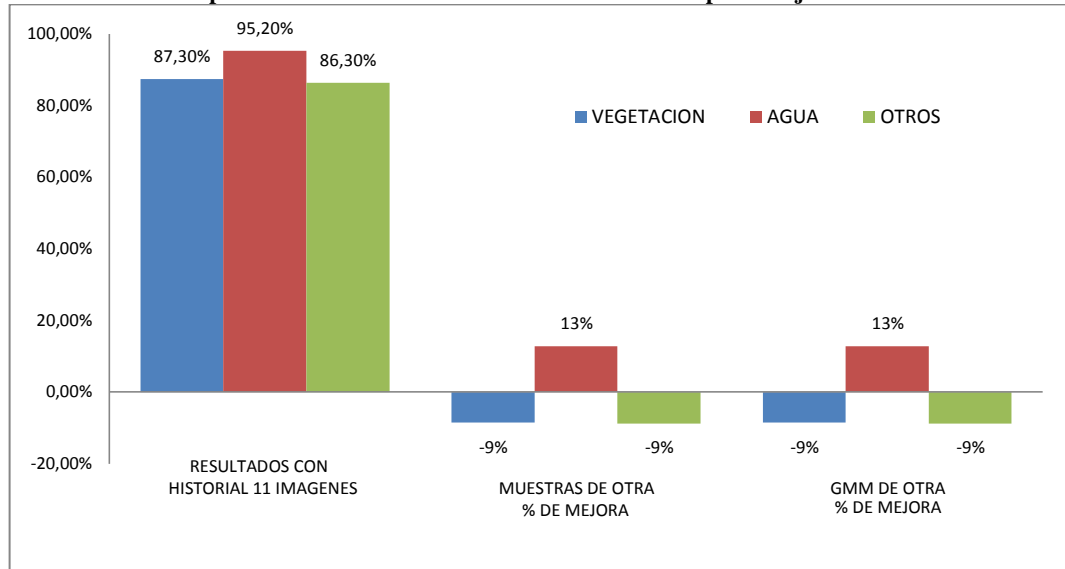
Finalmente se observa que con el algoritmo de segmentación semántica los resultados favorables son superiores o cercanos al 80% requeridos para 25 y 50 muestras y al final con todos los datos de entrenamiento disponibles superan el 85%, de esta forma se comprueba que los resultados pueden ser mejorados siempre que se tengan más ejemplos de las clases.

6.2.2 Análisis de resultados casos de segmentación

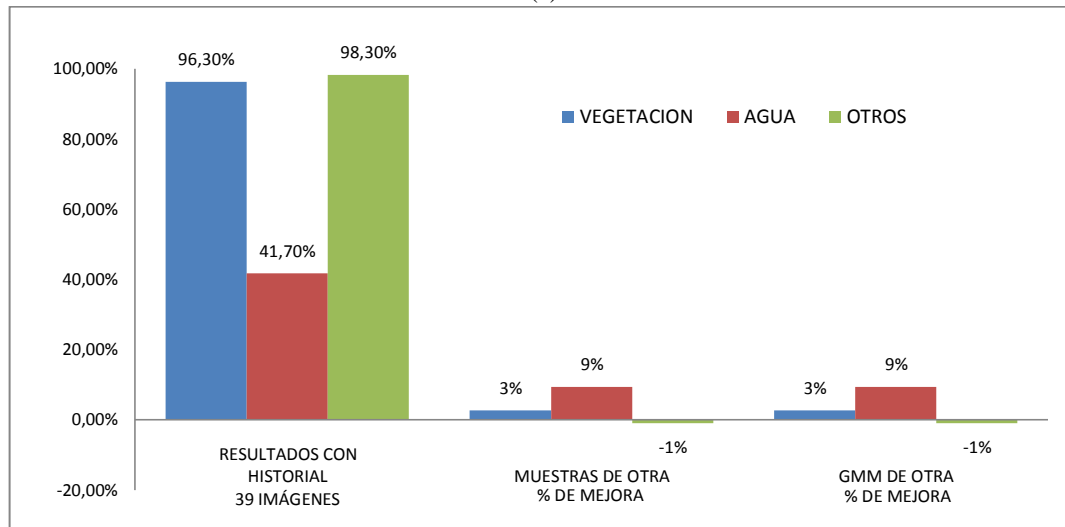
Revisando el método para segmentar imágenes a partir de muestras que se van añadiendo para el aprendizaje de la máquina, se hace un análisis que sirva para determinar que tanto se pueden mejorar los resultados. Con muestras

provenientes de varias imágenes aleatorias diferentes de sí misma, se observa que la calidad de la segmentación mejora considerablemente con respecto a la obtención de estos mismos resultados explicados en los numerales 5.4.6 y 5.4.7, la figura 22 muestra los resultados con las imágenes 011M.tif y 012M.tif.

Figura 22. Análisis de implementación del método adicionando datos de aprendizaje



(a)



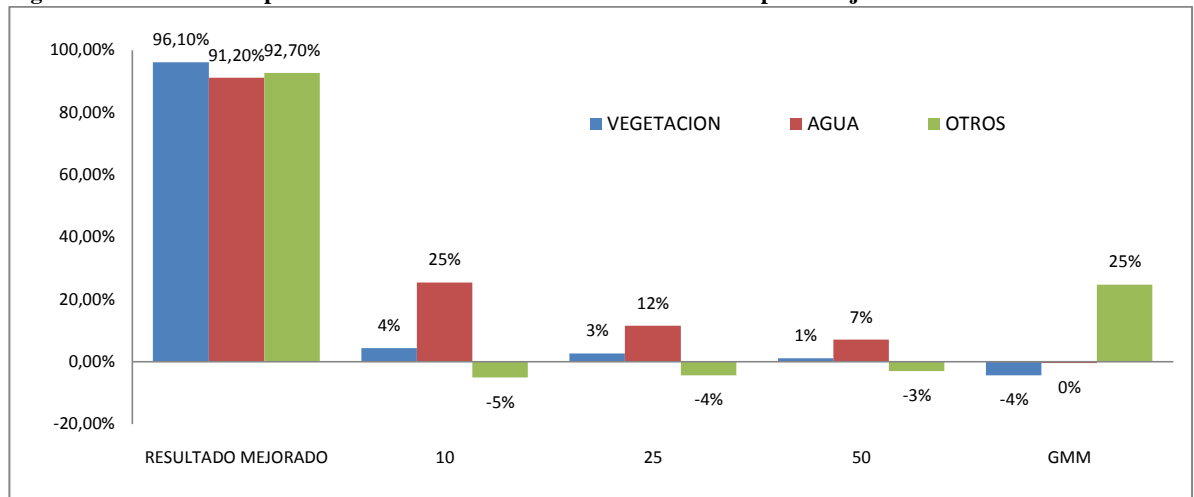
(b)

a) Resultados de segmentación con muestras de 10x10 de 11 imágenes aleatorias diferentes a la imagen segmentada 011M.tif comparado con dos métodos que toma información de otra imagen similar seleccionada por el usuario y con información de los clusters generados desde otra imagen con mezcla de Gaussianas. b) Resultados de segmentación con muestras de 10x10 de 39 imágenes aleatorias diferentes a la imagen segmentada 012M.tif, comparado con dos métodos que toma información de otra imagen similar seleccionada por el usuario y con información de los clusters generados desde otra imagen con mezcla de Gaussianas

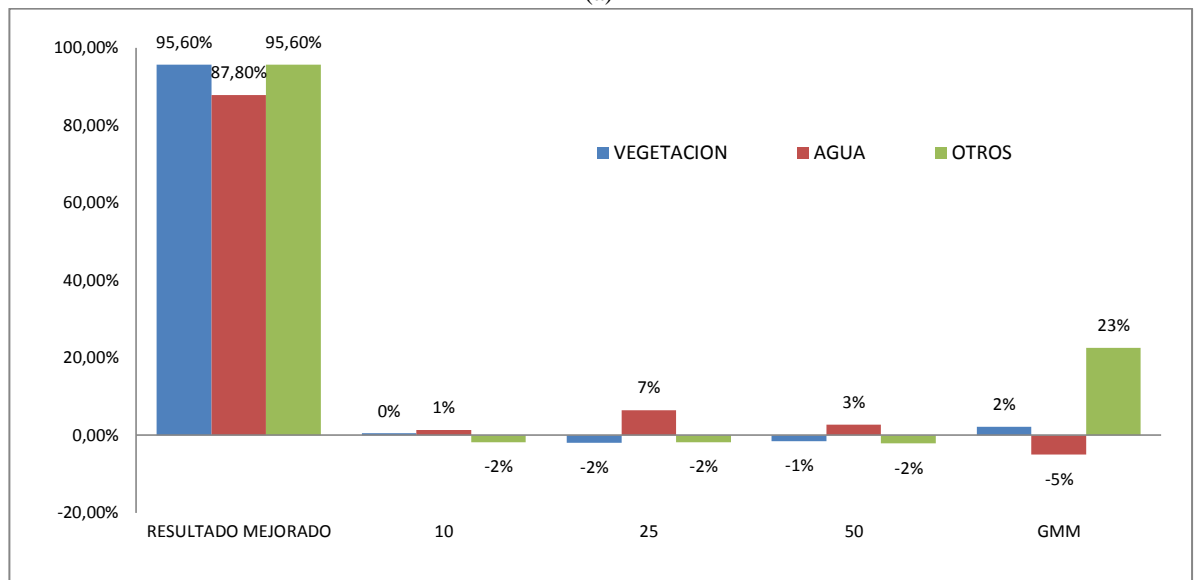
Para la imagen 011M.tif, los resultados ya están por encima del porcentaje de acierto en la clasificación de los píxeles, pero muestra un aumento considerable

en la precisión de clasificación para la clase Agua. En la imagen 012M.tif, también se observan mejores resultados en esta misma clase y en la clase vegetación, sin embargo, la clase agua está siendo mal clasificada más del 50% de los píxeles, luego la solución que se da para el método de segmentación semántica propuesto es tomar algunas muestras adicionales de la imagen misma para mejorar la calidad, los resultados se muestran en la figura 23.

Figura 23. Análisis de implementación del método adicionando datos de aprendizaje



(a)



(b)

a) Resultados de segmentación con muestras de 10x10 de la misma imagen a segmentar 011M.tif e información de 11 imágenes aleatorias comparado con el método de obtener 10, 25 y 50 muestras de la misma imagen y segmentación por pixel con el método de mezcla de Gaussianas. b) Resultados de segmentación con muestras de 10x10 de la misma imagen a segmentar 012M.tif e información de 39 imágenes aleatorias comparado con el método de obtener 10, 25 y 50 muestras de la misma imagen y segmentación por pixel con el método de mezcla de Gaussianas.

7. CONCLUSIONES

- Los resultados demuestran que el clasificador funciona con baja probabilidad de error sacando pocos datos de entrenamiento. Por otro lado el algoritmo consume pocos recursos computacionales, máxime teniendo en cuenta que se transformaron 928.800 datos de color con características en sus componentes RGB, en 150 muestras con 21 características obtenidas de su respuesta a los descriptores de covarianza en el entrenamiento y luego aplicándolos en la clasificación. Los resultados son superiores al 80% en segmentación correcta de las clases haciendo que al algoritmo aplicado la metodología de segmentación sea relevante y funcional para el usuario final.
- La técnica es aplicable y eficiente para la segmentación de imágenes de paisajes tropicales, además permite integrar otras características que ayudan a diferenciar más clases y con mayor precisión.
- Una ventaja que ofrece el algoritmo de segmentación semántica es que siempre se pueden mejorar los resultados en la medida que se tenga mayor número de muestras de entrenamiento dando la posibilidad de obtener resultados más exactos de los objetos contenidos en la imagen. Esta posibilidad no es aprovechable en los algoritmos de segmentación por píxel.
- Una desventaja encontrada en el desarrollo del algoritmo es que resulta muy dependiente de la disponibilidad del experto para recopilar datos para el análisis de error, razón por la cual no fue posible usar la totalidad de las imágenes disponibles en la base de datos para las pruebas y por otro lado algunas imágenes no cumplían los requisitos suficientes para las pruebas. De la misma forma se evidenció la dificultad para obtener imágenes para hacer otro tipo de análisis con sensores como el IR, ya que existe un grado de confidencialidad que no pudo superarse para obtenerlas, limitando el cumplimiento de los objetivos propuestos.

8. RECOMENDACIONES Y TRABAJOS FUTUROS

La base fundamental del algoritmo aplicado, consistió en la presentación de los datos en forma de sigma points y entrenados con SVM, es posible profundizar en el método si se emplea con otros métodos de entrenamiento de máquina y evaluando su desempeño.

El estudio de la segmentación semántica es aplicable a varios tipos de imágenes obtenidas con otros sensores, sin embargo, se debe evaluar el modo de presentación de los datos, para un ambiente supervisado. Este conocimiento es útil para mejorar el análisis y evaluación de imágenes cartográficas que en un futuro permitan mejorar la identificación de objetos extraños en imágenes de otros paisajes colombianos como el desierto, montaña, zonas costeras y marítimas donde son importantes estos estudios.

Para mejorar el análisis, velocidad de procesamiento y facilidad de uso por parte del usuario, es posible mejorar el algoritmo adicionando funciones semiautomáticas aprovechando las ventajas de presentar los datos a través de sigma points. Estos generan un grupo de números (en este caso 21) que relacionan el comportamiento de los componentes RGB de un grupo de píxeles, en el caso de las clases agua y vegetación parte de los sigma points se encuentran en intervalos muy pequeños ya sea por ejemplo porque la vegetación siempre es verde en distintas tonalidades o el agua siempre es plana sin importar su color lo que permite generar clústeres a través de mecanismos no supervisados como la mezcla de gaussianas para encontrar a priori la clase de un cuadro de píxeles seleccionado.

9. REFERENCIAS BIBLIOGRÁFICAS

- [1] Arbeláez, P.; Hariharan, B.; Gu, C.; Gupta, S.; Bourdev, L.; Malik, J.; University of California, Berkeley, *Semantic Segmentation using Regions and Parts*, Junio 2012, https://people.eecs.berkeley.edu/~sgupta/pdf/ahggbm_cvpr2012.pdf.
- [2] Avendaño, J.; Parra, J.; Bayona, J.; *Segmentación y clasificación de imágenes SAR en zonas de inundación en Colombia, una herramienta computacional para prevención de desastres*, Universidad Antonio Nariño – Revista de Facultades de Ingeniería, vol. 4, No. 8, pp 24-38; enero – Junio de 2014.
- [3] Carmona E.; Tutorial sobre maquinas vectores soporte (SVM), Universidad Nacional de Educación a Distancia UNED, 11 de Julio de 2014, p.1. [http://www.ia.uned.es/~ejcarmona/publicaciones/\[2013-Carmona\]%20SVM.pdf](http://www.ia.uned.es/~ejcarmona/publicaciones/[2013-Carmona]%20SVM.pdf)
- [4] Castilla G.; *Tesis Doctoral: Object-oriented analysis of remote sensing images for land cover mapping: conceptual foundations and a segmentation method to derive a baseline partition for classification*. Universidad Politécnica de Madrid Escuela Técnica Superior de Ingenieros de Montes, 2003.
- [5] Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B.; *The Cityscapes Dataset for Semantic Urban Scene Understanding*, Abril 2016, <https://arxiv.org/pdf/1604.01685.pdf>.
- [6] Guevara E.; Sánchez O. A.; “Segmentación de imágenes a color mediante técnicas de agrupamiento de datos empleando los algoritmos k-means y c-means”, Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas-IPN. México, D. F., Boletín No. 51, Nov de 2015
- [7] Girshick R.; Donahue J.; Darrell T.; Malik J.; “Rich feature hierarchies for accurate object detection and semantic segmentation”, CVPR2014, provided by Computer Vision Foundations, 2014.
- [8] Herrero A.; “Algoritmos para la estimación de modelos de mezclas gaussianas (GMM)” Escuela técnica superior de Ingenieros Industriales y de Telecomunicación, Universidad de Cantabria, Julio – 2015.
- [9] Julier, S., Uhlmann, J.K.: A general method for approximating nonlinear transformations of probability distributions. Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford (1996)

- [10] Kluckner S., Mauthner T., Roth P., Bischof H., Semantic classification in aerial imagery by integrating appearance and height information, Institute for Computer Graphics and Vision, Graz University of Technology, Austria, 2009. <http://www.icg.tugraz.at>.
- [11] P. Ponce, "Inteligencia Artificial con Aplicaciones a la Ingeniería" (1ª Edición) Alfaomega, México: ISBN: 978-607-7854-83-8. 2010.
- [12] Pacheco, C.; Del Valle, H.; Delrieux, C.; *Revisión de algoritmos de segmentación desde una perspectiva orientada al sensado remoto de imágenes ópticas y de radar*, Conference Paper, Agosto 2014. <http://www.researchgate.net/publication/269631987>.
- [13] Tuzel, O., Porikli, F., Meer, P.: Learning on lie groups for invariant detection and tracking, Mitsubishi Electric Research Laboratories, TR2008-031 July 2008, CVPR (2008). <https://pdfs.semanticscholar.org>
- [14] Ventura, C.; La segmentación semántica y sus benchmarks, [en línea], Universidad abierta de Cataluña, May 26, 2016, <http://informatica.blogs.uoc.edu/2016/05/26/la-segmentacion-semantica-y-sus-benchmarks/>.

10. ANEXOS

Adjunto pdf con la totalidad de las pruebas