

Clasificación y evaluación de una imagen satelital utilizando los métodos de redes neuronales artificiales y segmentación espacial

Satellite Image Classification and Evaluation using Neural Networks and Segmentation methods

Alejandra Gómez Sarmiento¹, Ingrid Rocío Buitrago Arévalo², Germán Torrijos Cadena³

Para citar este artículo: Gómez Sarmiento, A. Buitrago Arévalo, I.R. & Torrijos Cadena, G. (2014) Clasificación y evaluación de una imagen satelital utilizando los métodos de redes neuronales artificiales y segmentación espacial, Bogotá, Colombia. Revista de Topografía Azimut, PP 59-67

Fecha de Recepción: 20 de febrero de 2014

Fecha de Aceptación: 31 de agosto de 2014

Resumen

Los métodos de clasificación de imágenes satelitales han evolucionado y hoy en día se adoptan como métodos de investigación y avance en el uso de información satelital para provecho de diferentes áreas de la ingeniería, la aplicación de los métodos de clasificación de asignación de clases y el perceptrón multicapa (MLP). Estos, empleados en las áreas de segmentación y redes neuronales artificiales (RNA), se han implementado con el propósito de hacer una identificación e interpretación más clara de los diferentes tipos de coberturas en las zonas de interés. Los resultados obtenidos permitieron concluir el análisis de una serie de falencias y ventajas de cada uno de los métodos aplicado a una imagen obtenida para un departamento de Colombia.

Palabras clave: asignación de clases, perceptrón multicapa, segmentación, redes neuronales artificiales.

Abstract

Methods of classification of satellite images evolved nowadays concentrating itself as research and progress methods in the use of satellite information for the benefit of different areas of engineering, the application of class assignments and Multilayer Perceptron (MLP) methods, both employed in the areas of Segmentation and Artificial Neural Networks (ANN) have been implemented in order to make clearer the different types of coverage in areas of interest identification and interpretation. These results allowed us to conclude the analysis of a number of shortcomings and advantages of each of the methods applied to an image obtained for a department of Colombia.

Keywords: allocation of classes, multilayer perceptron, segmentation, artificial neural networks.

¹Ingeniera topográfica. alejandragomezarmiento@gmail.com

²Ingeniera topográfica. ing.ingridbuitragoud@gmail.com

³Ingeniero catastral y geodesta. germantorrijos@gmail.com

INTRODUCCIÓN

Actualmente existen diferentes tipos de plataformas contenidas de sensores dedicados a la captura y obtención de información de la superficie terrestre, una muestra de ello es la serie de satélites Landsat construidos y puestos en órbita por Estados Unidos para la observación en alta resolución de la Tierra. Hoy en día existe una lista de sensores con diferentes enfoques: el Landsat TM5 se dedica a la captura y producción de cartografía temática⁴, al igual que su predecesor, el *Thematic Mapper*, un equipo óptico electrónico de barrido multiespectral⁵. El incremento de la resolución espacial⁶ y espectral que ofrece este sensor facilita una mayor precisión para la cartografía temática (Chuvienco, 1990).

Lo anterior se utiliza con el fin de suministrar información útil y de pronosticar posibles fenómenos naturales, obtenidos por medio de métodos de clasificación y datos específicos de la imagen, al tener en cuenta el nivel de detalle de los datos. Hoy se emplean técnicas de extracción y agrupamiento de píxeles que presenten similitud espectral, así se generan agrupamientos por clases de píxeles con atributos en común o *clusters*, a esto se le denomina *clasificación de una imagen*. Este método se constituye en la distribución lógica de la información espacial para una interpretación más precisa de los datos.

Entre los métodos de clasificación empleados en la actualidad, existen los estadísticos, no estadísticos y otros, que emplean la inteligencia artificial. En este caso se utilizan los de segmentación espacial y redes neuronales artificiales, con los cuales se logra una comparación analítica y descriptiva de las clasificaciones por los métodos de asignación de clases y del modelo *perceptrón multicapa*. Todo este análisis se hace con el fin de determinar el método más eficiente

para la extracción y clasificación correcta de información, ya que se pretende que la precisión de la información clasificada para las clases a determinar sea lo más compatible posible con la realidad en el departamento del Vichada.

LA SEGMENTACIÓN: MULTIRRESOLUCIÓN

El proceso de separar la imagen en objetos se conoce como *segmentación*. Esta divide una imagen en objetos primarios que forman la base para clasificar el resto de la imagen. Es una técnica utilizada para analizar las imágenes satelitales, que se ha desarrollado hace relativamente poco tiempo, en comparación con el análisis tradicional de la imagen basada en píxeles (Burnett y Blaschke, 2003).

Aunque esta clasificación, al igual que la convencional, también se basa en información del píxeles (valores espectrales de cada píxel), adiciona un análisis basado en la información de un conjunto de píxeles similares llamados *objetos*. Estos objetos de la imagen son los grupos de píxeles que son similares entre sí sobre la base de una medida de las propiedades espectrales, es decir, color, tamaño, forma y textura, así como el contexto de un barrio que rodean los píxeles (Rodríguez, 2011).

El segmento se ha convertido en la unidad elemental para la simplificación de información en una imagen satelital, haciendo uso de los parámetros espectrales y espaciales para la clasificación de imágenes en teledetección. La *segmentación multirresolución* crea objetos mediante un algoritmo iterativo, por el cual se agrupan los objetos (a partir de píxeles individuales) hasta que se alcanza un umbral que representa la varianza del objeto superior. El umbral de varianza (parámetro de escala) se pondera

⁴Incluye los mapas de temas geográficos específicos orientados a propósitos específicos.

⁵Ocupa una franja de terreno recibiendo respuesta de varias bandas del espectro electromagnético.

⁶Número y ancho de canales que detecta el sensor.

con parámetros de forma (con separación de los parámetros de forma y compacidad) para reducir al mínimo los bordes fractales de los objetos. Entre los criterios más importantes a tener en cuenta para la extracción y organización de la información en la segmentación se emplean la *escala*, el *color*, la *suavidad* y la *compacidad*.

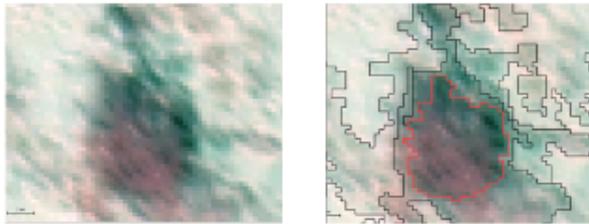


Figura 1: Segmentación espacial en una imagen

Fuente: Rodríguez (2011)

REDES NEURONALES ARTIFICIALES: EL PERCEPTRÓN MULTICAPA

Las redes neuronales artificiales surgieron de la necesidad de modelar la misma forma de procesamiento de la información en sistemas nerviosos biológicos y el cerebro humano (Izaurieta y Saavedra, s.f.). De esta forma, el funcionamiento del cerebro humano no es igual al de un computador digital convencional, ya que emplea un sistema altamente complejo no lineal paralelo basado en la resolución y ejecución de varias tareas al mismo tiempo. La arquitectura de una *red neuronal artificial* es una estructura contenida de un número de nodos interconectados con procesamiento automático, basado en el funcionamiento neurobiológico de los animales. Esta arquitectura genera una respuesta de salida conocida como *estímulo o excitación de la neurona*. En inteligencia artificial es frecuente referirse a ellas como *redes de neuronas o redes neuronales*. Cada neurona posee una característica de entrada/salida e

implementa una computación local o función. La salida de cualquier unidad de información está determinada por su característica de entrada/salida, su interconexión de otras unidades y posiblemente de sus entradas externas. Sin embargo, es posible un trabajo a mano, la red desarrolla usualmente una funcionalidad general a través de una o más formas de entrenamiento (Schalkoff, 1997).

Las características más importantes de las redes neuronales les atribuyen similitudes con respecto al funcionamiento y procesamiento del cerebro humano, donde se destacan el aprendizaje adaptativo, la tolerancia a fallos que le permiten capacidades para retener información aun sufriendo una destrucción parcial de la red, la autoorganización de información, y su fácil adaptabilidad a diferentes tareas en campos diversos.

El modelo del perceptrón fue introducido en el 1950 por Ronsblat, quien adoptó la idea por Mc-Culloch y Pitt; pero además introdujo un método automático para encontrar los valores de los pesos y umbrales. Rosenblat fue el pionero para introducir el aprendizaje de tipo supervisado para entrenar la red. Esta RNA es un modelo neuronal con propagación hacia adelante que se caracteriza por la organización en capas de celdas disjuntas, de modo que la salida de una neurona o su respuesta no es enviada a las neuronas de la misma capa o de la capa anterior

La neurona biológica o artificial puede estar activa o inactiva por medio de una *función de activación* por una función de entrada menos el umbral. En la actualidad acuden como funciones de activación a la función lineal, sigmoidea y tangente hiperbólica.

EXTRACCIÓN DE INFORMACIÓN Y CLASIFICACIÓN DE LA IMAGEN

Composición RGB y llaves de interpretación

La combinación RGB (4,5,3) fue escogida por su eficiente capacidad de discriminar tipos de vegetación, debido a que contiene una gran sensibilidad al follaje que se identifica en color rojo; los bosques más densos (primitivos) reflejan un tono rojo más oscuro y opaco; en cambio la vegetación inundable se detalla en una tonalidad rojo oscuro o vino tinto, que se encuentra ubicada en las partes más cercanas a los cuerpos de agua meándricos o lagos; a su vez el agua se observa en un color similar al azul oscuro o negro debido a la absorción de energía junto con la sedimentación en los ríos meándricos que afloran un color blanco o de tono claro, puesto que refleja la luz solar.

Las selvas se identifican con un rango de colores que comprenden rojos o rojos muy oscuros, el color cian representa las zonas de suelos áridos, desnudos, senderos y áreas de establecimiento urbano que reflejan la energía solar.

Cabe resaltar que la imagen multispectral ofrece una interpretación más precisa de los componentes de la imagen debido a que las características de un canal permiten obtener información acerca de una clase que sea escasa de estudiar en una banda contraria, así se podrá asegurar que la obtención de datos de la clasificación será recogida en casi su totalidad puesto que se relacionan las tres bandas espectrales:

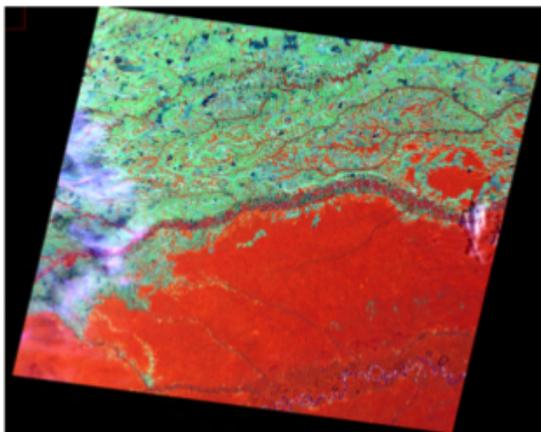


Figura 2: Composición RGB a Falso color mejorado de la Imagen

roja (R), infrarroja cercana (IRC), infrarroja media (IRM).

Segmentación multirresolución y aplicación de la diferencia espectral

El proceso de segmentación de la imagen se obtuvo con la definición de los pesos asignados a cada una de las canales en la combinación RGB, en este caso se asignó un peso doble (2) a la banda infrarroja cercana, porque a partir de esta se puede hacer una discriminación más profunda de vegetación en la imagen.

Se definieron parámetros de escala, forma y compacidad, asegurando que la creación de objetos se hiciera de la forma más precisa entre los píxeles. Luego se determinó un parámetro de refinamiento espectral en información de la imagen.

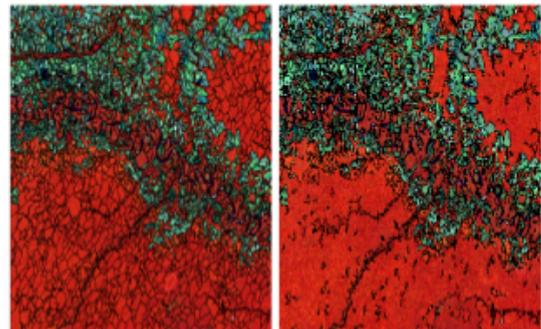


Figura 3: Segmentación multirresolución y aplicación de la diferencia espectral

Entrenamiento y arquitectura de la red

La topología correcta para la red neuronal se escogió a partir de un constante entrenamiento e iteración de valores para cada parámetro, con el objetivo de mejorar la respuesta deseada en la red neuronal; para este caso solamente se tomó una capa oculta en el procesamiento con cinco neuronas para lograr eficiencia en la información de salida evitando cometer un alto error y generar alguna confusión a la red neuronal en su entrenamiento. En la figura 4 se muestra el

esquema de arquitectura seleccionado para la retropropagación del error.

El número de neuronas dispuesto para la red de entrada fueron cuatro (4), lo que coincidió con las bandas empleadas para la discriminación de vegetación como lo son: la banda roja (R), banda infrarroja cercana (IRC) y la banda infrarroja media (IRM) además de una entrada de NDVI aplicado a la imagen para el aumento en la precisión de la diferenciación de los tipos de vegetación presentes en la imagen.

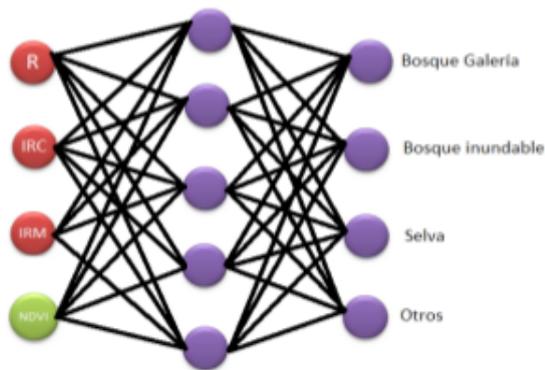


Figura 4: Arquitectura de la red neuronal artificial

En la estructura de la red se definieron cuatro (4) unidades de salida como el *Bosque Galería*, el *Bosque Inundable*, *Selva* y una clase denominada *Otros* para incluir todo aquello que no sea de interés para esta clasificación. Cabe resaltar que la mayoría de escenarios presentes en esta clase está constituida por sabanas, cuerpos de agua profundos, ríos meándricos como el Vichada y el Guaviare, entre otros.

La cantidad de neuronas escogidas para el buen desempeño de la red fue cinco (5) con el propósito de mejorar el procesamiento de la red pero sin recargarla con una cantidad excesiva de unidades. La técnica habitual para obtener solución al entrenamiento de la red neuronal se denomina *retropropagación*, que consiste en modificar el procesamiento de la red neuronal por medio de la propagación

hacia adelante y hacia atrás, en este caso se tomaron muestras puras de entrenamiento para cada clase y se definieron a lo largo del proceso los pesos asignados para la obtención del error y su respectiva propagación hacia atrás, comparando los resultados obtenidos por la red y los esperados por el usuario.

El entrenamiento comenzó con la introducción de entradas (canales y NDVI) a la red con sus respectivas muestras de entrenamiento tomadas de la imagen además de la definición de parámetros arbitrarios de *tasa de aprendizaje* y *momento*. El RMS permitido se estableció en 0.01 debido a que es un valor fiable para el entrenamiento de la red, y su intervalo de asignación oscila entre 0.000 a 0.05 y el obtenido cumplió con el umbral concluido en el valor de 0.001.

Los valores ideales para la clasificación de la imagen para la tasa de aprendizaje y momento fueron de 0.4 y 0.99, dando así una tasa de aprendizaje final de 0.000, la cual muestra que la red aprendió en su totalidad y la imagen fue clasificada de manera más precisa después de un entrenamiento comprendido por 133 épocas y 4000 iteraciones.

Se establecieron desde un principio valores arbitrarios para la *tasa de aprendizaje* y *factor momento* con datos iniciales de 0.2 y 0.5, buscando que la red aprendiera de forma más eficiente a lo largo de las épocas y de esta forma se cumplió con la clasificación esperada, incluyendo las cuatro clases en el resultado final, además de discriminar los elementos importantes en la clase de *Otros* donde no existía vegetación de interés. La gráfica lograda a través del entrenamiento con 4000 iteraciones para una tasa final de aprendizaje de 0.000, se muestra en la figura 5.

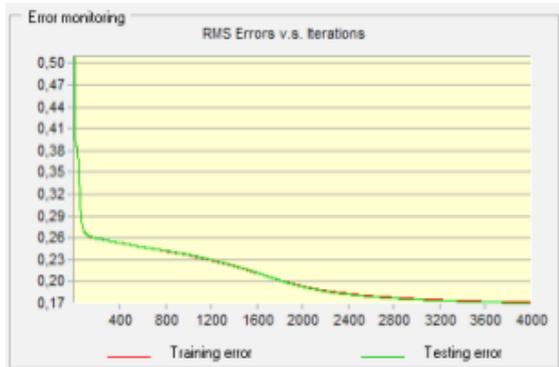


Figura 5: Función de activación sigmoidea para 4000 Iteraciones

RESULTADOS Y DISCUSIÓN

Clasificación por asignación de clases

La asignación de clases se configuró, previo reconocimiento de distintos tipos de cobertura vegetal en la imagen satelital, de manera que las clases elegidas fueron *Bosque Galería*, *Bosque Inundable*, *Selva* y *Otros*.

Para cada una de las clases escogidas se empleó un intervalo de valores que oscilaban entre el rango de -1 a 1, correspondientes a la condición del NDVI. En la clasificación para la clase *Otros* se definió un umbral para el $NDVI \leq 0.7$ debido a que en ella los valores de NDs son más variables ya que por contener una gran cantidad de sabanas, lagos, cuerpos de agua profundos y con caudal alto –además de incluir la sedimentación en los ríos Vichada y Guaviare a causa de su forma meándrica– no son parte de la densa vegetación y bosques de diferentes tipos a discriminar. La tabla 1 muestra los umbrales para cada categoría.

Clase	Umbral para NDVI
Bosque Inundable	$0.7 < NDVI \leq 0.754$
Bosque Galeria	$0.754 < NDVI \leq 0.82$
Selva	$0.82 < NDVI \leq 1$
Otros	$NDVI \leq 0.82$

Tabla 1: Tabla de umbrales para la asignación de clases

La clasificación por asignación de clases mediante umbrales refleja exactitud debido a que no hubo información redundante ni mal clasificada; cada clase coincidió adecuadamente con su llave de interpretación como se puede visualizar en la figura 6.

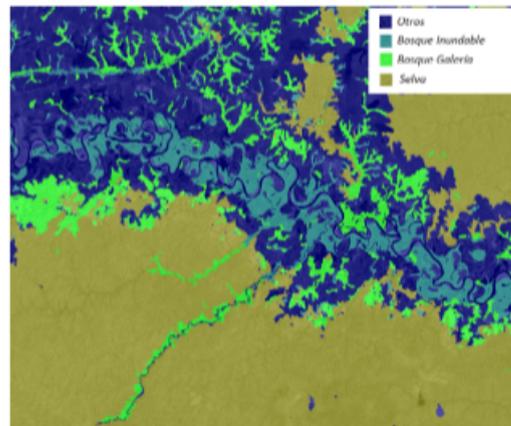


Figura 6: Clasificación de la imagen satelital por asignación de clases

Clasificación por el método de MLP

La siguiente es la clasificación por el clasificador MLP donde se asignaron los píxeles de la imagen a las clases: *Bosque Inundable*, *Bosque Galeria*, *Selva* y *Otros*, como se observa en la figura 7.

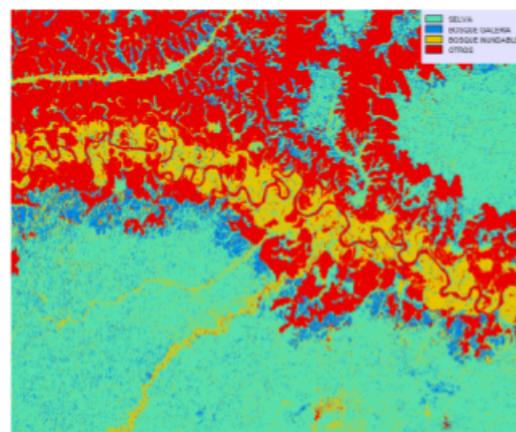


Figura 7: Clasificación de la imagen satelital por el método MLP

Debido a que la función logarítmica oscila entre los valores de 0 a 1, la función se desarrolló entre los datos de 0.17 a 0.50, generando un solo mínimo local y dando un resultado para los errores de entrenamiento y de control de 0.1702 y 0.1692 (figura 5). Por esta razón, la precisión obtenida de esta época fue de 93.32% para la clasificación de una imagen Landsat.

Evaluación de las clasificaciones

En este apartado se utilizaron dos procesos de evaluación, siendo el primero la matriz de confusión y el segundo el índice de concordancia Kappa, calculados a partir de datos estadísticos y por lo cual se recurrió a un diseño de muestreo aleatorio estratificado por clase para la verificación de la evaluación de las clasificaciones.

Estadísticos: clasificación por asignación de clases

Se puede detallar en la *matriz de confusión* (tabla 2) que los errores de comisión y omisión fueron completamente nulos, por tanto la exactitud del productor y del usuario fue exitosa al 100%, ya que ningún píxel quedó sin agrupar y todos los segmentos se clasificaron en la imagen satelital.

Estadísticos: clasificación por MLP

La matriz de confusión aplicada demuestra en este caso, la clase con menor cantidad de píxeles confundidos es la clase *Otros* (tabla 5). Las exactitudes del productor y usuario asimismo como el error de omisión como de comisión para la clase de *Bosque Galería* fue la más baja de todas, y el porcentaje más alto se evidenció en la clase de *Selva*, debido a que el total de píxeles para esta clase es mayor que el de *Bosque Galería*, por esta razón pudieron ser incluidos erróneamente en esta última.

Los datos estadísticos calculados para llegar al porcentaje del índice Kappa, como resultado de la evaluación de la clasificación, fueron calculados directamente de los

resultados de los totales por clase por muestra escogida, por lo cual se produjo un alto porcentaje en la fiabilidad real a causa del error estándar del muestreo que concluyó en un valor menor de 0.01.

Matriz de Confusión						
Clasificación	Bosque Inundable	Bosque Galería	Selva	Total	Error Comisión (%)	Exactitud del Usuario (%)
Bosque Inundable	55	0	0	55	0,00	100,00
Bosque Galería	0	93	0	93	0,00	100,00
Selva	0	0	52	52	0,00	100,00
Total	55	93	52	200		
Error Comisión (%)	0,00	0,00	0,00			
Exactitud del Productor (%)	100,00	100,00	100,00			

Fuente: Elaboración propia

Tabla 2: Matriz de confusión para la clasificación por asignación de clases

Datos Estadísticos	
Sumatoria Diagonal	200
Fiabilidad Global	1
Error proporcional Total	0
Abscisa de la curva normal	1,96
Error estandar del muestreo	0
Fiabilidad Real	100,00 %
	Intervalo de fiabilidad
	100,00 %

Fuente Elaboración Propia

Tabla 3: Datos estadísticos para el cálculo del coeficiente Kappa en la clasificación de asignación de clases

Índice Kappa	
n	200
Sumatoria diagonal	200
Producto de totales	14378
	100,00 %

Fuente Elaboración propia

Tabla 4: Índice Kappa para la clasificación por asignación de clases

Matriz de Confusión							
Clasificación	Selva	Bosque Galería	Bosque Inundable	Otros	Total	Error Comisión (%)	Exactitud del Usuario (%)
Selva	10429	1345	671	0	12445	16,20	83,80
Bosque Galería	434	2590	102	0	3126	17,15	82,85
Bosque Inundable	567	354	5908	8	6837	13,59	86,41
Otros	0	2	0	27588	27590		
Total	11430	4291	6681	27596	49998		
Error Comisión (%)	8,76	39,64	11,57				
Exactitud del Productor (%)	91,24	60,36	88,43				

Fuente: Elaboración propia

Tabla 5: Matriz de confusión para la clasificación por el MLP

Datos Estadísticos	
Sumatoria Diagonal	46515
Fiabilidad Global	0,930337213
Error Proporcional Total	0,069662787
Abscisa de la curva normal	1,96
Error estandar del muestreo	0,00113853
Fiabilidad Real	93,26 %
	Intervalo de fiabilidad
	92,81 %

Fuente Elaboracion Propia

Tabla 6: Datos estadísticos para el cálculo del índice Kappa en la clasificación por MLP

El coeficiente Kappa determinó el grado de precisión para la clasificación previa con un valor de 0.88 determinado *casi perfecto* (tabla 7), teniendo en cuenta los rangos de evaluación del índice de concordancia.

Índice Kappa		
n	49998	
Sumatoria diagonal	46515	88,67 %
Producto de totales	962711653	

Fuente Elaboracion propia

Tabla 7: Índice Kappa para la clasificación por MLP

Partiendo de una muestra de 49,998 píxeles en la clasificación por MLP empleado en las RNA, se encontraron 3,475 píxeles mal clasificados y un alto error de omisión y comisión en la clase *Bosque Galería* con una exactitud del productor de 60.36 % y 82.85 % en cada uno de los errores. De esta forma la clase con menores errores y con exactitud más alta fue la clase *Otros* (en este caso el programa tomó esta categoría como una unidad de salida requerida para el entrenamiento de la red).

Por otro lado, se hizo el mismo diseño y extracción en la clasificación por asignación de clases para el cálculo de la matriz de errores.

En consecuencia, se distingue de forma notoria que en la matriz de confusión no hubo ningún segmento mal clasificado, lo que correspondió a una asignación exacta de cada objeto a su respectiva categoría, y con lo cual se logró una precisión del 100 %.

En las tablas 4 y 7 se encuentran las muestras correspondientes para la RNA y la segmentación; así, se asignaron 49,998 píxeles de muestra para el MLP y 200 segmentos para la asignación de clases. Estadísticamente se registra una diferencia considerable en los valores de la muestra debido a que el segmento simplifica la información y contiene una agrupación de píxeles con semejanza espectral y espacial, por consiguiente se evita la generación de información redundante. En la sumatoria de la diagonal de la matriz de confusión para el método MLP se obtuvo

un valor de 3,475 píxeles que fueron mal clasificados, por tanto, el índice Kappa tuvo un valor de 88.67% menos preciso, en comparación con la asignación de clases donde el índice de concordancia dio un 100 % de precisión, debido a que la sumatoria de la diagonal coincide con el tamaño de la muestra.

CONCLUSIONES

El perceptrón multicapa es un modelo que emplea el clasificador MLP y su respectivo aprendizaje se denomina *retropropagación* y fue utilizado en este trabajo para el entrenamiento de la red. Este método presenta desventajas en cuanto al tiempo de convergencia debido a que tarda en aprender la red neuronal artificial, pues se desconoce una estructura general de inicio para entrenar dicha red.

La asignación de clases utilizó umbrales para la delimitación de las categorías entre intervalos cerrados de ND de NDVI con el objetivo de no dejar ningún píxel sin clasificar o en una clase no correspondiente; esto agrega precisión, rapidez, facilidad de interpretación, agilidad al proceso de asignación, disminución del tiempo y veracidad en el producto, como se comprobó en la clasificación por asignación de clases. De igual manera se confirmó en el índice de concordancia Kappa para esta clasificación, pues se obtuvo el valor más alto: 100 % de precisión.

Comparando el conocimiento y la experiencia en estas dos tecnologías se observa que es más eficiente el segmento en vez del píxel, ya que el primero involucra un aprovechamiento de la resolución espacial al producir dos veces más la cantidad de parámetros a considerar al momento de generar un segmento. Así, la combinación de un tratamiento básico de la imagen a falso color mejorado y sumando la aplicación de la tecnología mencionada con uso de umbrales en la asignación de clases supera los mecanismos tradicionales de extracción de muestras en una imagen porque no se

supedita a la fiabilidad del productor sino que se le otorga más oportunidad de decisión al usuario. Esto se evidencia en el empleo de umbrales con intervalos cerrados y no con una muestra que puede presentar errores desde su obtención.

Por último, se evidenció el resultado de las dos clasificaciones a través del índice de concordancia donde este tuvo exactitud en el método de asignación de clases mediante la segmentación multirresolución con respecto al perceptrón multicapa de las redes neuronales artificiales, que logró un grado de precisión confiable.

REFERENCIAS BIBLIOGRÁFICAS

- Burnett, C. & Blaschke, T. (2003). A multi-scale segmentation / object relationship modelling methodology for landscape analysis. *Ecological Modelling*, 168,233-249.
- Chuvieco, E. (1990). *Fundamentos de teledetección espacial*. Madrid: Ediciones Rialp S.A.
- Izaurieta, F. & Saavedra, C. (s.f.). *Redes neuronales artificiales*. Concepción: Universidad de Concepción.
- Rodriguez M., A. (13 de octubre de 2011). *Geoinformación: para entender los recursos de la Tierra*. Recuperado el 24 de febrero de 2013, de: <http://geoinformacion.wordpress.com/2011/10/13/clasificacion-basada-en-objetos-object-based-image-analysis-obia/>
- Schalkoff, R.J. (1997). *Artificial Neural Networks*. Nueva York: McGraw-Hill.