

CLASIFICACIÓN DE IMÁGENES MULTIESPECTRALES UTILIZANDO REDES NEURONALES SUPERVISADAS

Diana Cristina Ruiz-Alvarez

División Física Aplicada, Depto. de Ciencias de la Computación, CICESE

Km. 107 Carret. Tijuana-Ensenada, Ensenada B.C., México

E-mail: druiz@cicese.mx

RESUMEN

Se presenta la implementación de un sistema de clasificación basado en Redes Neuronales, así como los resultados obtenidos de su aplicación a una imagen multiespectral adquirida del sensor Landsat MSS, correspondiente a la zona deltaica del Río Colorado. La clasificación fue realizada utilizando una red neuronal de retropropagación que fue entrenada para reconocer hasta quince diferentes tipos de terreno. Los resultados obtenidos durante el entrenamiento, muestran hasta once clases de terreno reconocidas con base en cuatro características dadas como entrada para cada tipo de terreno.

INTRODUCCIÓN

La interpretación de imágenes de satélite es ejecutada con dos tipos de métodos (Richards, 1986): el análisis cuantitativo que se basa en el procesamiento digital de imágenes utilizando computadoras, y el análisis cualitativo donde la inspección visual y la información contextual, como es la proximidad a la región de estudio, es tomada en cuenta por el interpretador para definir la forma y el tipo del terreno.

Un elemento importante en la interpretación de datos remotos es la identificación de terrenos. Dicha interpretación es una tarea adecuada para los métodos cuantitativos de análisis, ya que la computadora es capaz de utilizar todos los componentes espectrales obtenidos del sensor, mientras que el análisis humano no lo permite.

La clasificación de tipos de terreno a partir de imágenes de satélite es de gran utilidad en estudios a nivel regional, así como en áreas donde se requiere de un monitoreo constante. La importancia de este tipo de análisis radica en su aplicación a estudios de impacto ambiental, planeación regional y de áreas urbanas, así como en la administración de recursos naturales.

La clasificación supervisada es el procedimiento más frecuentemente utilizado para el análisis cuantitativo de datos en el sensado remoto de imágenes. Existen algunos algoritmos por medio de los cuales se puede llevar a cabo esta clasificación; algunos de ellos se basan en modelos de distribución probabilísticos para las clases, otros en donde las clases son divididas en regiones y otros más, como lo son las redes neuronales, se basan en la estadística. Independientemente del modelo utilizado para este tipo de clasificación, existen 5 pasos esenciales a seguir (Richards, 1986):

- Decidir el conjunto de tipos de terreno dentro del cual la imagen va a ser segmentada. Por ejemplo: agua, vegetación, etc.

- Escoger los píxeles representativos para cada uno de los tipos de terreno elegido. Estos píxeles formarán el conjunto de datos de entrenamiento.
- Usar los datos de entrenamiento para determinar los parámetros del clasificador a utilizar. Al conjunto de parámetros para una clase también se le llama firma de la clase.
- Clasificar cada punto de la imagen dentro de alguno de los tipos de cubierta elegidos.
- Producir tablas o mapas temáticos los cuales suman los resultados obtenidos.

Durante los últimos años se ha propuesto el empleo de Redes Neuronales Artificiales como una herramienta alternativa a los métodos de clasificación tradicionales (clasificador de máxima similitud o el de distancias mínimas). Los métodos tradicionales de clasificación se basan en la suposición de que la función de densidad de probabilidad de los datos se comporta como una distribución de Gauss. Esta dependencia de un modelo probabilístico representa su principal limitante. Las Redes Neuronales representan una alternativa a este tipo de clasificadores, ya que al no basarse en un modelo probabilístico pueden ser usadas en un contexto más amplio.

Se han realizado trabajos de clasificación utilizando las Redes Neuronales (Heermann y Khazenie, 1992, Bischof *et al.*, 1992), ya que éstas poseen la habilidad para generalizar, hacer suposiciones acerca de la estadística de los datos de entrada y la capacidad para formar límites de decisión no lineales (Rumelhart *et al.*, 1986). Heermann y Khazenie (1992) utilizaron la red de retropropagación para clasificar una imagen multiespectral y concluyeron que este es un método factible para clasificar volúmenes de datos muy grandes.

En este trabajo se presenta la implementación de un clasificador basado en una red neuronal de retropropagación,

donde ésta es entrenada para reconocer hasta quince tipos de terreno. En la primera parte de este trabajo se presenta la red de retropropagación, su arquitectura y funcionamiento. En la siguiente sección se detalla la implementación del proceso a seguir para realizar la clasificación mediante redes neuronales, la formación de los patrones de entrenamiento y la elección de la estructura de la red. Por último, se presentan los resultados obtenidos en las pruebas realizadas.

REDES NEURONALES

Una red neuronal esta compuesta por un conjunto de elementos de procesamiento llamados neuronas, las cuales se comunican a través de interconexiones de peso o fuerza variable.

La red de retropropagación está compuesta por al menos tres capas de neuronas, una capa de entrada que reciba los estímulos del exterior, una capa de salida que genere una respuesta y una o más capas intermedias. Las neuronas de cada capa se comunican con cada una de las neuronas de la capa superior inmediata, sin existir comunicación dentro de la misma capa (Figura 1).

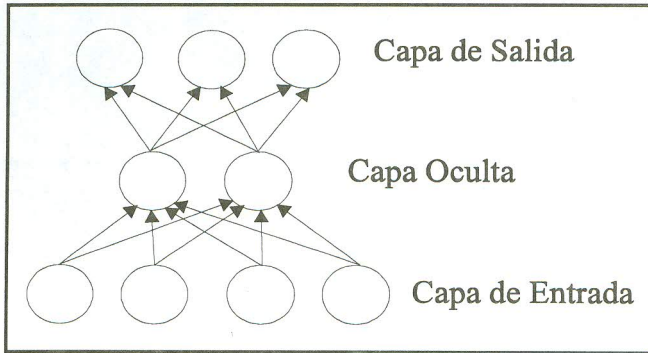


Figura 1. Estructura de la red de retropropagación.

Cada elemento en la red recibe una entrada de la capa inferior, la cual procesa para generar una salida, que es transmitida a las neuronas de la capa superior, que realizarán el mismo proceso (Figura 2).

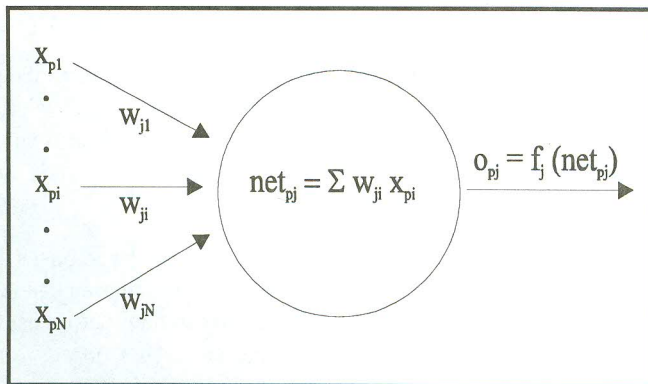


Figura 2. Operaciones de la neurona.

La red aprende un patrón de entrada-salida por medio de un ciclo de propagación-adaptación. En la primera etapa, un

patrón X_p formado por N componentes, denotado por X_{pi} donde $i = 1, n$ es dado como estímulo a la capa de entrada.

La capa de entrada tiene como función propagar el patrón a las demás capas. La entrada total a una neurona en una capa superior es calculada mediante la sumatoria

$$net_{pj} = \sum W_{ji} X_{pi}, \quad (1)$$

donde W_{ji} representa el peso de la conexión de la neurona i a la neurona j , X_{pi} es el componente del patrón de entrenamiento cuando se trata de la primera capa intermedia, o la salida de la neurona i en la capa anterior, si se trata de la capa de salida.

El valor de salida de una neurona, llamado valor de activación, es calculado mediante una función. Para la red de retropropagación generalmente se utiliza la función sigmoideal:

$$f(net) = (1 + e^{-net})^{-1} \quad (2)$$

por lo que la salida de la neurona es igual a

$$y_p = f(net). \quad (3)$$

Este proceso es repetido capa por capa hasta obtener una salida de la red. Entonces se calcula el error generado mediante

$$E_p = \frac{1}{2} \sum (y_p - o_p)^2, \quad (4)$$

donde o_p es la salida esperada para el patrón p .

La segunda parte del proceso es la adaptación o retropropagación, donde los pesos de las interconexiones son modificados con base al error, permitiendo que la red aprenda, por lo que a este proceso se le llama entrenamiento.

La regla de aprendizaje utilizada por la red es llamada Regla Delta Generalizada. Esta se basa en la modificación continua de los pesos de las conexiones para reducir la diferencia (delta) entre la salida obtenida y la salida esperada. La actualización de los pesos es determinada por la Regla del Gradiente Descendente. Los pesos son modificados en una cantidad proporcional a la primera derivada del error con respecto a los pesos

$$\Delta W_{ji} \propto - \frac{\partial E_p}{\partial W_{ji}} \quad (5)$$

Por lo tanto la modificación de los pesos de las conexiones estará dada por

$$W_{ji}(t+1) = W_{ji}(t) + \Delta p W_{ji} + \mu \Delta p W_{ji}(t-1), \quad (6)$$

donde

$$\Delta p W_{ji} = \eta \delta_{pj} X_{pi}. \quad (7)$$

Dentro de la actualización de los pesos se considera la tasa de aprendizaje de la red (η) y un término llamado momentum (μ). La tasa de aprendizaje determina la velocidad de aprendizaje de la red, de esta forma, al asignar una tasa de aprendizaje baja, la red tomará más tiempo en completar el entrenamiento.

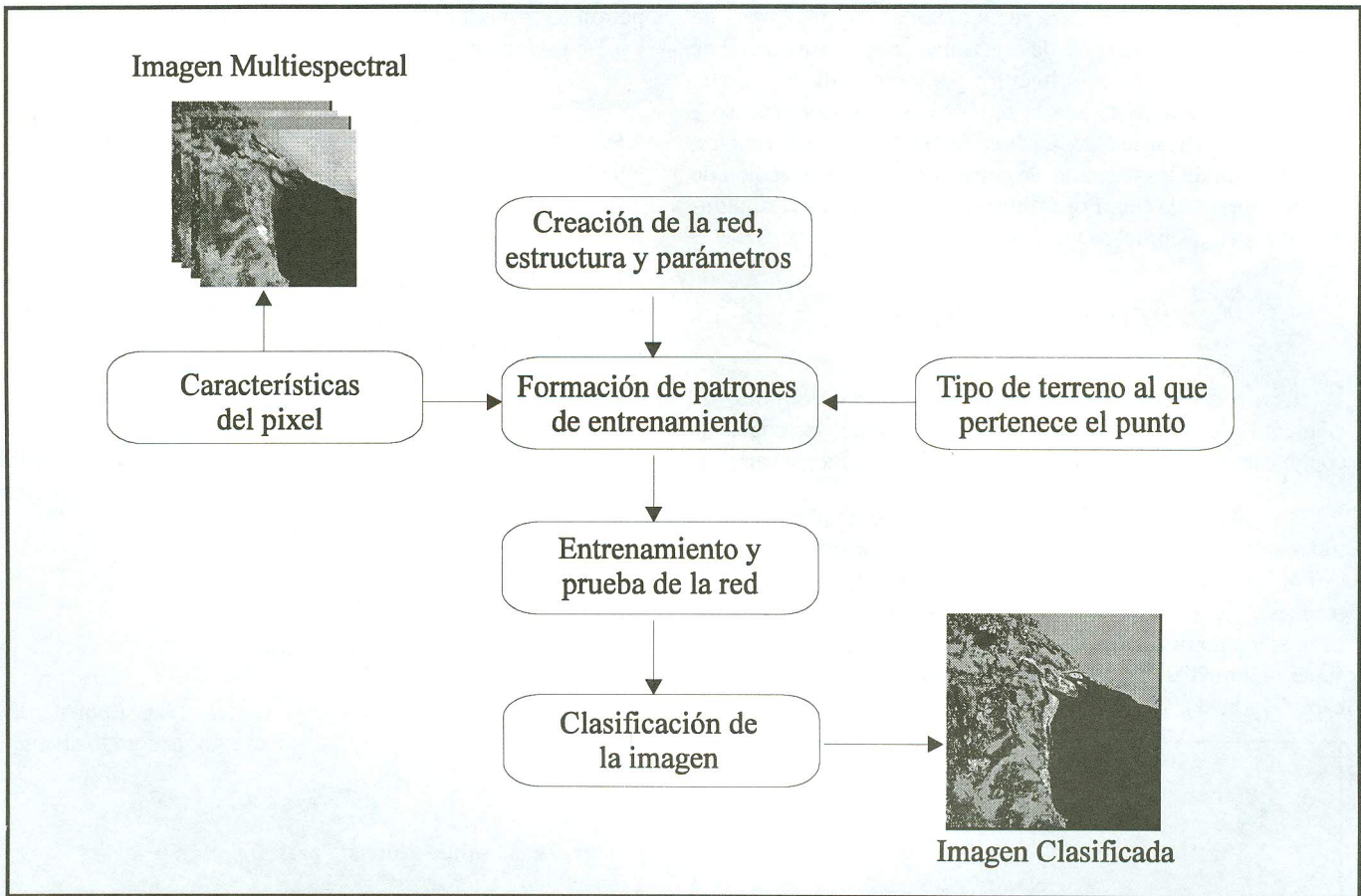


Figura 3. Proceso de clasificación.

El parámetro denominado momentum tiene como objetivo ayudar a mantener la dirección a seguir para realizar el cambio de los pesos. Estos dos parámetros generalmente son positivos y menores que 1.

IMPLEMENTACIÓN

El proceso de clasificación de imágenes multiespectrales utilizando redes neuronales (Favela *et al.*, 1994) está definido por una serie de pasos similares a los que se siguen en cualquier método tradicional (Figura 3).

Primeramente se debe definir la estructura de la red que será utilizada. El número de nodos en las capas de entrada y salida se determina de acuerdo a la aplicación, ya que la capa de entrada recibe las características del punto que se desea clasificar, firma espectral, altitud, pendiente, etc., y la capa de salida representará el tipo de terreno para el punto seleccionado. Dado que la imagen consta de cuatro bandas, la capa de entrada fue diseñada con cuatro neuronas, donde cada neurona recibe el valor del pixel en cada una de las bandas. La capa de salida se formó con 15 neuronas, que son el número de clases. Estos valores se conservaron durante todas las pruebas. Para determinar la estructura interna de la red, dado que no existe una metodología para elegir el número óptimo de capas

intermedias u ocultas y el número de nodos en cada una de ellas, es necesario experimentar para encontrar la estructura que mejores resultados proporcione. Para el número de capas ocultas y la cantidad de nodos en cada una de ellas se realizaron varias pruebas, variándose entre una y tres capas intermedias y entre 7 y 12 neuronas para cada capa. Se utilizaron entonces diferentes estructuras de red para realizar las clasificaciones; la primera de ellas constó de una capa interna con 8 neuronas, la segunda estructura utilizada fue de dos capas internas con 8 y 12 neuronas, respectivamente. Por último, se utilizó una red de tres capas ocultas con 7, 10 y 12 neuronas, respectivamente.

Para determinar los parámetros de la red (tasa de aprendizaje y momentum), es necesario realizar pruebas con distintos valores y seleccionar el óptimo. La tasa de aprendizaje se varió entre 0.3 y 0.8, y el momentum se varió entre 0.1 y 0.5, para cada una de las estructuras.

El siguiente paso en el proceso de clasificación es la formación del conjunto de patrones que se presentarán a la red durante el entrenamiento (patrones de entrenamiento). Estos proporcionan ejemplos a la red acerca de los puntos que encontrará durante la clasificación. Cada patrón de entrenamiento consta de dos vectores, el primero de ellos es el de entrada, cuyo número de componentes es determinado por la cantidad de características elegidas para el punto, este se formó

tomando los valores del pixel en cada una de las cuatro bandas, por lo que se tiene un vector con cuatro componentes. El segundo vector es el de salida, donde el número de tipos de terrenos a reconocer determina de cuantos componentes constará, por lo que se tienen vectores de 15 componentes. Para representar los diferentes tipos de terrenos se asigna un valor de 1 a un solo nodo y un cero a los demás. Dependiendo de la posición del uno en el vector, será la clase representada (Figura 4). De esta manera, se formaron dos diferentes conjuntos de patrones con los cuales se entrenó a la red. En el primero de ellos se tienen tres patrones para cada clase, mientras que en el segundo se tomaron cinco patrones por clase.

Cada componente del patrón fue normalizado con el fin de

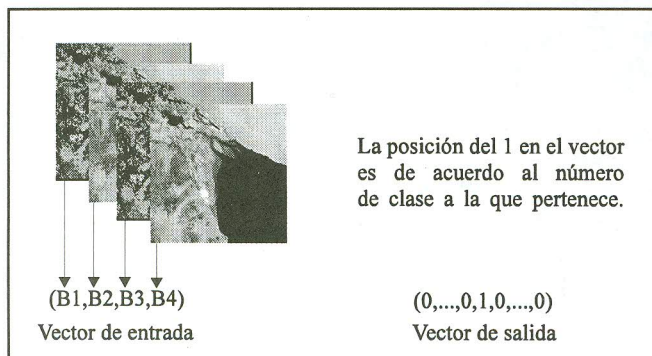


Figura 4. Formación de patrones de entrenamiento.

tener valores dentro del intervalo (0,1), requerido para la operación de la red. Para formar el conjunto de patrones de entrenamiento se realizó una selección de puntos en la imagen a partir de un mapa temático de la zona (Torres-Rodríguez, 1994), de modo que éstos representaran a las diferentes clases de terreno.

Construidos los patrones, éstos son presentados a la red un número de veces determinado durante la fase de entrenamiento, lo que permite que sean “aprendidos” y puedan ser reconocidos durante la clasificación. A continuación, se realiza una prueba de la red para determinar si el entrenamiento fue satisfactorio o es necesario repetirlo con un número de iteraciones mayor o variando los parámetros. Por último, se realiza la clasificación de la imagen completa de la cual se obtiene una nueva imagen donde se muestran las clases encontradas.

El entrenamiento utilizado consistió de 5000 y 10000 ciclos de presentación para cada patrón.

RESULTADOS

La imagen utilizada en las pruebas proviene del satélite LANDSAT 3 Sensor Multi Spectral Scanner (MSS), capturada el 16 de enero de 1980 y corresponde a la región sur del delta del Río Colorado (Figura 5). Esta consta de cuatro bandas espectrales (4, 5, 6 y 7), donde cada pixel puede tener un valor entre 0 y 127, el cual representa la reflectancia de la cobertura

terrestre para cada banda espectral.

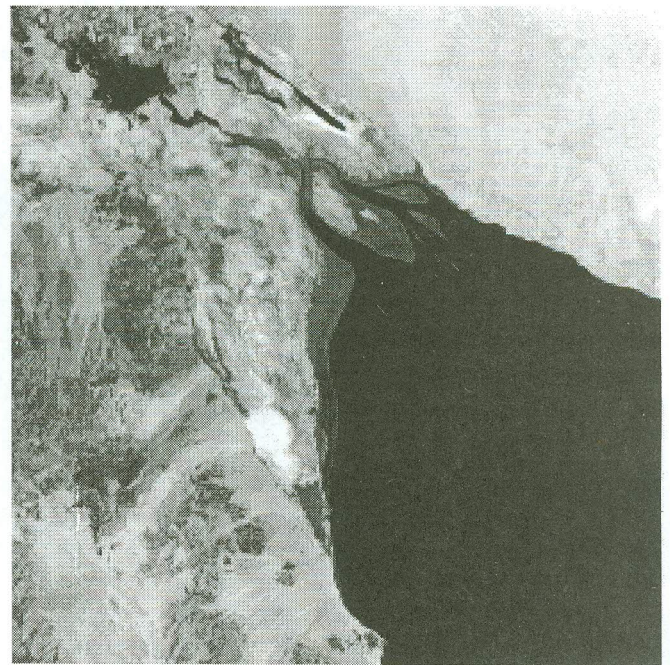


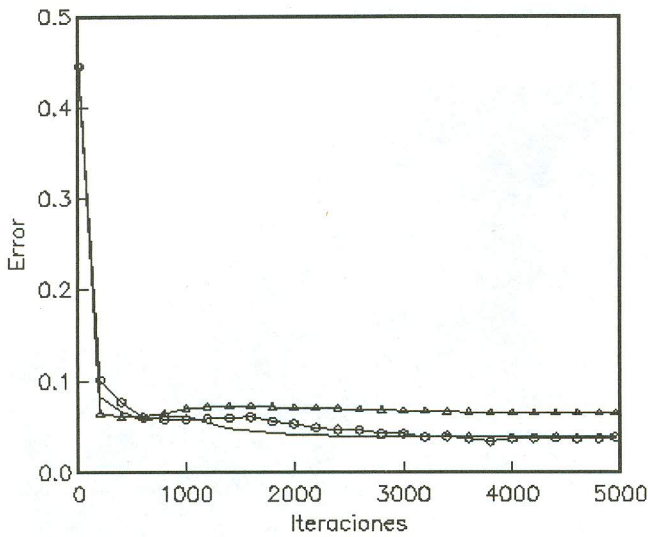
Figura 5. Imagen utilizada para la clasificación (banda 7).

Se han realizado diferentes aplicaciones en los últimos años para la clasificación de imágenes satelitales utilizando redes neuronales (Heermann y Khazenie, 1992, Bischof *et al.*, 1992). Los resultados obtenidos en ellas fueron comparados con métodos tradicionales mostrando que el empleo de las redes neuronales es factible en la clasificación de imágenes multiespectrales. En estas aplicaciones las redes fueron utilizadas para clasificar un número pequeño de tipos de terreno, generalmente entre 3 y 7 clases, con base en un mayor número de características de entrada. En la aplicación aquí presentada, la red es utilizada para clasificar hasta 15 tipos de terreno con base en cuatro características dadas como entrada.

Diversas pruebas fueron realizadas, variando tanto el número de capas y nodos en ellas, como los diferentes parámetros de la red. Así mismo, se varió el número de patrones de entrenamiento por tipo de terreno y la cantidad de ciclos de presentación de los mismos.

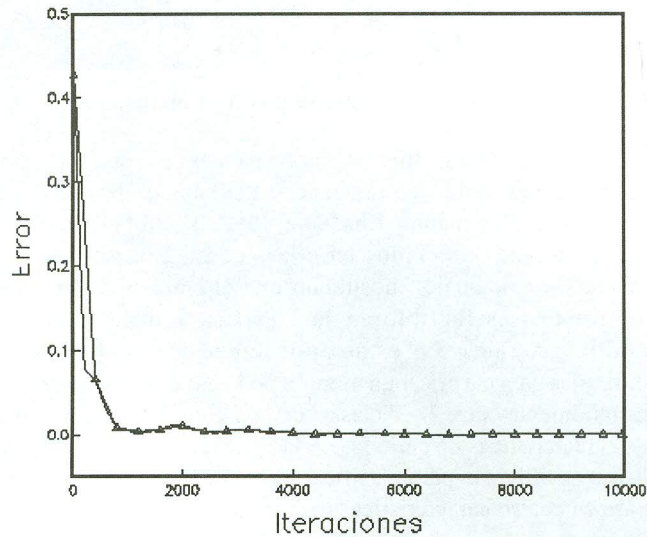
La variación del número de capas de la red, influyó, tanto en el error generado como en el tiempo de entrenamiento. Contrario a lo que se esperaba, al agregar demasiadas capas internas a la red se incrementa el error generado y el tiempo de entrenamiento; esto último se debe a que deben realizarse más operaciones con cada capa que se agrega. Se determinó que la mejor estructura fue aquella con dos capas intermedias. Los resultados de estas pruebas se muestran en la Figura 6.

En lo que respecta a la cantidad de presentaciones para cada uno de los patrones de entrenamiento, ésta solo influye en el tiempo de aprendizaje, ya que en determinado momento, la disminución del error es casi nula, por lo que no es recomendable



○ 1 capa — 2 capas ▲ 3 capas

Figura 6. Error como función del número de capas internas.

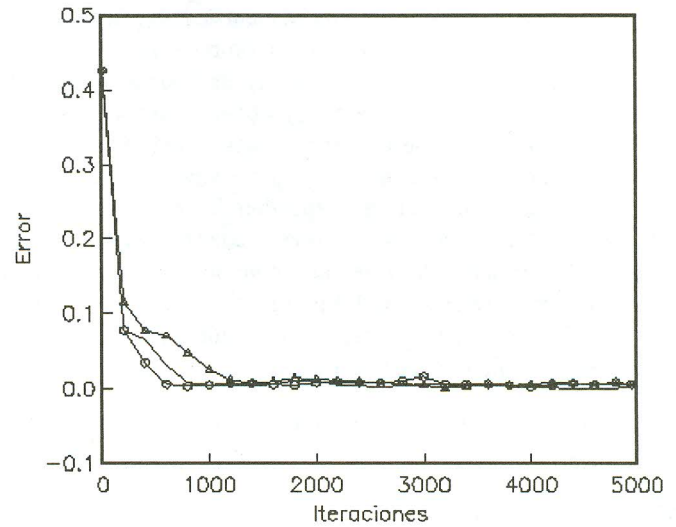


— 5000 iteraciones ▲ 10000 iteraciones

Figura 7. Error como función del número de presentaciones de los patrones de entrenamiento.

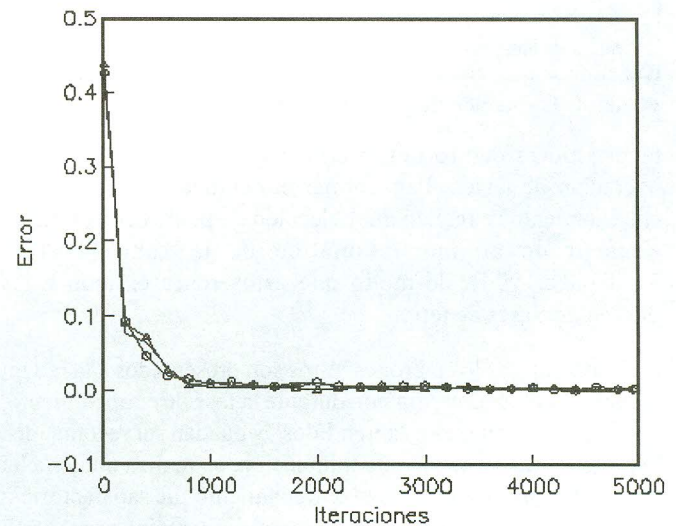
tener un número de iteraciones muy grande. Los resultados obtenidos en las pruebas son presentados en la Figura 7.

Con respecto a los diferentes parámetros de la red, éstos se variaron para la mejor estructura de red, cada uno de ellos se varió dejando a los demás fijos. El cambio de la tasa de aprendizaje determinó la velocidad con que la red aprende los patrones presentados. Al asignar un valor mayor para este parámetro el error en el entrenamiento fue mayor al principio, ya que la red trata de aprender más rápidamente cada patrón, en cambio, al asignar una tasa de aprendizaje menor, los patrones son aprendidos más lentamente, y por lo tanto el error es menor. Una comparación de entrenamiento con diferentes tasas de



○ $\eta=0.3$ — $\eta=0.5$ ▲ $\eta=0.8$

Figura 8. Error como función de la tasa de aprendizaje. aprendizaje es presentado en la Figura 8.



○ $\mu=0.1$ — $\mu=0.3$ ▲ $\mu=0.5$

Figura 9. Error como función del parámetro momentum.

El término momentum influyó en el descenso del error durante el entrenamiento. Así con un valor mayor para éste, se obtuvo una disminución del error más suave y sin tantas fluctuaciones. El cuanto al error final, la diferencia no fue considerable (Figura 9).

El número de patrones presentados para cada clase fue importante en la clasificación, ya que se proporcionan más ejemplos de las características de cada tipo de terreno a encontrar durante la clasificación. Así, en las pruebas realizadas se obtuvieron mejores resultados al tener un mayor número de patrones por clase.

Los mejores resultados fueron obtenidos con una red que consta de dos capas ocultas, con 8 y 12 neuronas en cada una de ellas. La red tiene una valor para la tasa de aprendizaje de 0.5 y para el momentum de 0.3. El conjunto de patrones de entrenamiento consta de 75 patrones, 5 por cada una de las 15 clases a reconocer. Estos patrones fueron presentados 5000 veces a la red, la tasa de error obtenida fue de 0.001.



Figura 10. Imagen clasificada.

Los resultados obtenidos de la clasificación son 11 clases reconocidas (Figura 10). Dentro de los tipo de terreno encontrados, de acuerdo con el mapa temático utilizado (Torres-Rodríguez, 1994), se identifican: planicie alta, planicie intramareal, planicie intermareal y planicie supramareal, zona de cultivo y zona de depresión; en la península de Baja California la zona de relieve, depósitos no consolidados y zona de erosión; en Sonora, el desierto; por último, el agua se consideró como un tipo distinto de terreno.

Si el lector esta interesado en conocer más a fondo los resultados obtenidos durante este trabajo, puede contactar con el autor del mismo.

AGRADECIMIENTOS

Un sincero agradecimiento a Jesús Favela Vara y a Jorge Torres Rodríguez por el apoyo y colaboración brindado durante la realización de este trabajo, así como a Alejandro Hinojosa Corona por la ayuda ofrecida durante la revisión del trabajo escrito y a Hugo Hidalgo por el arbitraje realizado.

REFERENCIAS

- Bischof, H., Schneider, W., and Pinz, A.J., 1992, Multispectral Classification of Landsat Images Using Neural Networks, IEEE Transactions on Geoscience and Remote Sensing, 30, 482-490 p.
- Favela, J., Torres, J., Hidalgo, H. and Granillo R., 1994, Load Balancing for Neural Network Classification of Remote Sensing Data in an Heterogeneous network of Workstation, Proc. of the 7th Intl. Conf. on Parallel and Distributed Processing.
- Heermann, P.D., and Khazenie, N., 1992, Classification of Multispectral Remote Sensing Data Using Back-Propagation Neural Network, IEEE Transactions on Geoscience and Remote Sensing, 30, 81-88 p.
- Richards, J., 1986, Remote Sensing Digital Image Analysis, Springer-Verlag, Berlin, Heidelberg, 281 pp.
- Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1996, Learning Internal Representation by error propagation. In. Rumelhard, D.E., and McClland, J.L. (Eds.), Parallel Distributed Processing: Exploration in the Microstructure of Cognition Volume 1: Foundations, MIT Press, 318-362 p.
- Torres-Rodríguez J., 1994, Aplicación de sensores remotos (Landsat MSS) para el reconocimiento de unidades geomorfológicas en la región Deltáica del Río Colorado, Tesis Maestría, Facultad de Ciencias Marinas, U.A.B.C., 117 pp.