

CLASIFICACIÓN NO SUPERVISADA DE IMÁGENES HIPERESPECTRALES UTILIZANDO UNA RED NEURONAL AUTO-ORGANIZATIVA

P. MARTÍNEZ, P. L. AGUILAR, A. PLAZA, M. LINAJE, J. C. PRECIADO y R. M. PÉREZ,

pablomar@unex.es

*Dpto. de Informática. Universidad de Extremadura.
Avda. de la Universidad s/n. Cáceres*

RESUMEN: Este trabajo realiza la clasificación no supervisada de una imagen hiperespectral de referencia denominada Indian Pines, mediante una red neuronal SOM. Los parámetros de la red (función de aprendizaje y función de vecindad) se ajustan en base a una batería de firmas espectrales de pruebas generada a partir de la propia imagen, fijando los valores que proporcionan mayor robustez y capacidad de discriminación. La exactitud en la clasificación proporcionada por la red se analiza a partir de la matriz confusión, y se establece la discusión en base al número de neuronas utilizado en la capa de salida, al número de iteraciones y la distancia inicial que controla el aprendizaje competitivo. Los resultados muestran una exactitud del productor media para las neuronas de salida cercana al 80% (60% considerando las no clasificadas) que disminuye al utilizar menor número de neuronas de salida.

ABSTRACT: This paper make the unsupervised classification of a hyperspectral image, known as Indian Pines, with a SOM Neural Network. The parameters of the net (learning function and neighbourhood function) are adjusted from a large battery of hyperspectral firms generated from the Indian Pines image, correcting them to obtain a stronger and more effective discrimination in the neural net. The results of the classification (the output of the net) are analysed through the confusion matrix, and the discussion is established about the number of neurons in the output layer, the number of iterations and the starting distance which control the competitive learning. The results show an accuracy about 80% for the output neurons (about 60% if we considerate the non-classificated pixels) which it's decremented if we use less output neurons.

Palabras clave: Análisis Hiperespectral, Redes Neuronales, SOM, Clasificación No Supervisada, Teledetección.

INTRODUCCIÓN

Los sensores hiperespectrales (s.h.) son instrumentos que combinan la disposición espacial de un sensor de imágenes con las capacidades analíticas de un espectro-radiómetro.

Para cada píxel de la imagen un s.h. produce un espectro completo que se conoce como firma hiperespectral (f.h.) del píxel. Una f.h. puede considerarse como un vector que almacena la radiancia de la zona de la cubierta correspondiente en varias bandas espectrales estrechas y contiguas (radiancia espectral). El estudio de la superficie terrestre mediante imágenes hiperespectrales, permite ampliar el rango de aplicaciones de la teledetección y mejorar la precisión de las metodologías convencionales.

Los algoritmos convencionales tienen dificultades para manejar los datos hiperespectrales por su alta dimensionalidad (Harsany, 1994); es necesario utilizar nuevos métodos.

Podemos enumerar las siguientes razones (Bruske, 1999) para considerar a las redes neuronales como un candidato idóneo para el tratamiento de datos hiperespectrales:

Robustez para manejar información contaminada por diversas fuentes de ruido

Paralelismo intrínseco, ofreciendo facilidades para tratar grandes cantidades de datos.

Parámetros ajustables para adaptar el funcionamiento de la red a cada problema concreto.

Método intuitivo que permite ocultar la complejidad del aparato matemático y manejarlo con una interfaz amigable.

Simplicidad. La mayor parte de los algoritmos neuronales son procesos iterativos simples que involucran operaciones matriz por vector.

De los diferentes modelos de computación neuronal, las redes neuronales auto-organizativas S.O.M., se caracterizan por establecer una correspondencia que preserva la topología (Kohonen, 1997), esta cualidad las hace especialmente interesantes para resolver problemas de clasificación en espacios N -dimensionales, y ha sido utilizada en estudios similares, (Aguilar, 2000).

METODOLOGÍA

La red SOM propuesta para la clasificación de imágenes hiperespectrales utiliza M neuronas de salida (una por cada clase que se prevé que tengan los patrones de aprendizaje) y N neuronas de entrada (una por cada banda del espectro del píxel), la descripción detallada del algoritmo de aprendizaje puede encontrarse en Kohonen (1997) y Aguilar (2000).

La clasificación llevada a cabo por la red depende de los siguientes parámetros:

Topología de la capa de salida

La elección más sencilla consiste en utilizar una capa unidimensional, con sus extremos conectados

Inicialización de pesos

Los valores iniciales para los prototipos (pesos) han sido seleccionados de forma aleatoria.

Función de vecindad

Utilizamos una función Gausiana:
 $\gamma(t,i,i^*)[x(n)] = \exp(-dist(i^* - i) / \sigma(t))$
 Donde $\sigma(t)$ denota la distancia de referencia.
 La función utilizada es $\sigma(t) = \sigma_0 / t$.

Función de aprendizaje

Este parámetro debe ser decreciente en el tiempo.
 La función utilizada es $\eta(t) = 1/t$

Imagen de Test

La imagen usada para validar la red neuronal propuesta, corresponde a una escena denominada Indian Pines (I.P.), que fue obtenidas por el sensor AVIRIS a bordo de un avión ER2 en Northern Indiana el 12 de Junio de 1992 con una resolución espacial de 17 m.

La imagen I.P. contiene 145x145 píxeles, cada uno con 200 bandas, para cada píxel la imagen almacena la radiancia del sensor sin correcciones atmosféricas.

Para reducir el tiempo de aprendizaje y prueba, utilizamos una subimagen I.P. que corresponde a una porción para la que disponemos de verdad terreno (g.t.).

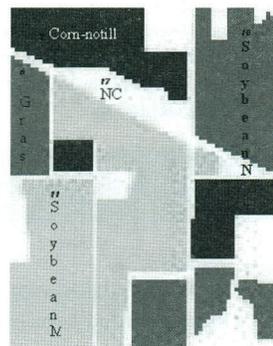


Figura 1. Imagen Verdad Terreno (g.t.)

La elección de esta imagen se debe a la dificultad de su clasificación, por la poco favorable relación señal ruido y el alto grado de correlación entre firmas.

La imagen completa contiene 17 clases (la clase 17 corresponde a píxeles no clasificados N.C.). La subimagen contiene 4 clases cuyo identificador aparece en la Figura 1, junto a su leyenda.

El conjunto de entrenamiento está formado por todas las firmas de la subimagen. Durante el proceso de aprendizaje recorremos todos los píxeles de la subimagen, a partir de un punto inicial aleatorio, distinto en cada iteración t . Una vez obtenidos los prototipos para cada neurona, clasificamos cada píxel y obtenemos la matriz de confusión.

Para analizar la exactitud de la clasificación proponemos una medida similar a la exactitud del productor (Chuvieco, 2000).

Como la red SOM trabaja como un algoritmo no supervisado, no existe una biyección entre clases - verdad terreno y los valores máximos no tienen que aparecer en la diagonal principal. Al ser mayor el número de neuronas que el de clases g.t., pueden participar varias neuronas de salida en una misma verdad terreno, sin que ello suponga una clasificación incorrecta. Los síntomas de una clasificación errónea se ponen de manifiesto cuando una o más columnas de la matriz de confusión presentan valores altos para varias clases g.t.

Para obtener una medida de la capacidad de una neurona i de la red, para clasificar la imagen, calculamos la proporción entre el valor máximo para una columna de la matriz de confusión $X_{max,i}$ y la suma de todos los valores de la misma columna, $X_{s_i} = \sum_j X_{j,i}$ tal y como aparece en la ecuación 1:

$$\text{Ecuación 1} \quad E_{ni} = 100 \frac{X_{max,i}}{X_{s_i}}$$

Este valor ofrece información sobre la capacidad de cada neurona para discriminar entre clases y puede promediarse para todas las neuronas, obteniendo una medida global de precisión de la clasificación realizada, que denominaremos "exactitud media".

RESULTADOS

En las clasificaciones se modificaron el número de neuronas de la capa de salida y el número de iteraciones.

Primera prueba (r100n16)

En la primera prueba se utilizaron 100 iteraciones y 16 neuronas. La figura 2a muestra la imagen clasificada.

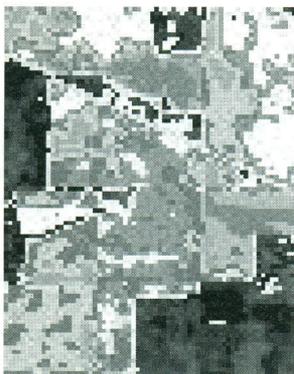


Figura 2a. Imagen Clasificada r100 n16.

Para discutir los resultados hemos de tener en cuenta que cada columna de la matriz de confusión en niveles de gris, (figura 2b), corresponde a una neurona de salida y que las situaciones más favorables aparecen en neuronas que se activan exclusivamente para una clase, discriminándola de las restantes. Gráficamente esta cualidad puede observarse en la g.t. 6, (hierba), de la figura 2b.

El solapamiento de columnas, como el que aparece entre soja (10 y 11) y maíz (2), indica una clasificación inadecuada que puede deberse a unos parámetros mal seleccionados, a una delimitación incorrecta de la verdad terreno, o a que el algoritmo no es adecuado para este tipo de clasificaciones.

Las neuronas de la red SOM que se activan para una cierta clase g.t. deben respetar la topología, este hecho se pone de manifiesto por la continuidad de las franjas en la matriz de confusión. La clasificación realizada para las variedades de soja (10 y 11) muestra valores aislados que nos indican que la red no ha respetado la topología de dichas clases, y puede deberse a la

separabilidad de las clases, especialmente cuando los máximos en dichas filas se cruzan.

Hay que señalar que la clase 17 se solapa con todas las otras ya que corresponde a los píxel no clasificados. La exactitud media de la clasificación fue del 60%, a pesar de incluir a dicha clase en nuestros cálculos, siendo del 80% sin considerar los puntos de la clase 17.

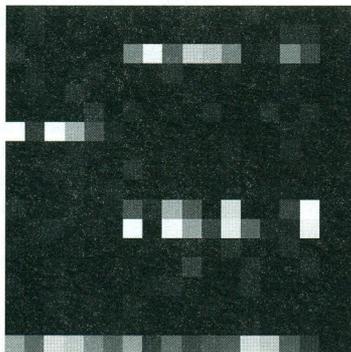


Figura 2b. Imagen Matriz de Confusión r100 n16.

Segunda prueba (r500n5)

En esta prueba se utilizaron 5 neuronas de salida y 500 iteraciones, manteniendo los restantes parámetros. La figura 3a muestra la imagen clasificada.

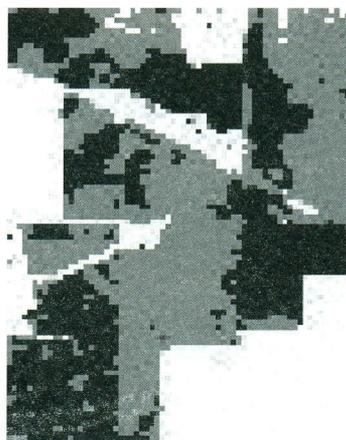


Figura 3a. Imagen Clasificada r500 n5.

En la imagen en niveles de gris, correspondiente a la matriz de confusión, (figura 3b), podemos observar que los resultados obtenidos en cuanto a exactitud de la clasificación son peores (exactitud media del 55%) ya

que las primeras neuronas se activan prácticamente por igual para las dos clases de soja (10 y 11) y maíz (2), discriminando correctamente la hierba (6), las clases aprendidas han preservado mejor que en la primera prueba la topología, desapareciendo las discontinuidades dentro de las clases.



Figura 3b. Imagen Matriz de Confusión r500 n5.

Tercera prueba (r500n10)

En esta prueba se utilizaron 10 neuronas de salida y 500 iteraciones, manteniendo los restantes parámetros. La imagen clasificada puede verse en la figura 4a.



Figura 4a. Imagen Clasificada r500 n10

Los resultados para este experimento muestran una mayor exactitud media (60%) debida al menor solapamiento de las clases, dicho solapamiento solamente se mantiene para las dos tipos de soja, lo que indica la dificultad que conlleva dicha separación. La utilización de un mayor número de neuronas hace que no se respete la topología, apareciendo varias "islas" para el maíz y la soja en la matriz de confusión (figura 4b).

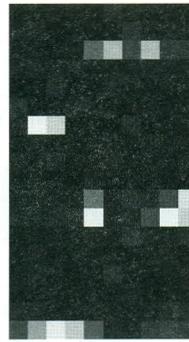


Figura 4b. Imagen Matriz de Confusión r500 n10.

Estos tres experimentos son representativos del conjunto de pruebas realizadas que presentan la misma tendencia.

CONCLUSIONES

La red neuronal utilizada forma clases usando un aprendizaje sin supervisión. Los resultados indican que al aumentar el número de neuronas, aumenta la exactitud media, pero no se respeta la topología de dichas clases. Disminuyendo el número de neuronas, la competencia entre ellas asegura una mayor coherencia, reduciéndose la exactitud. Para mantener el mismo nivel de competencia con un número de clases suficiente, se podría utilizar una distancia de referencia que disminuya suavemente, y con un valor inicial mayor, consiguiendo así que compitan las neuronas más alejadas.

BIBLIOGRAFÍA

- AGUILAR MATEOS, P. L., PÉREZ R. M., MARTÍNEZ P., HORMIGO A. & DÍAZ P. "Abundance extractions from AVIRIS image using a Self-Organizing neural network", Proc IX AVIRIS Workshop, pp 2000.
- BRUSKE, J. & MERÉNYI, E., 1999, Estimating the intrinsic dimensionality of Hyperspectral images. Proc. European Symposium on Artificial Neural Network, ESANN'99, Bruges, Belgium, 21-23 April, 1999, pp. 105-110
- CHUVIECO, E. "Fundamentos de Teledetección Espacial", Edt Rialp, 2000.
- HARSANYI & CHANG, "Hyperspectral image classification and dimensionality reduction", IEEE TGARS vol. 32, pp 779-785, 1994.
- KOHONEN, T. "Self-Organizing Maps (2nd. ed.)", Springer Series in Information Science. 1997.

AGRADECIMIENTOS

Proyecto financiado por el M.C. y T. (Ref. TIC 2000-0739-C04-03).

Este trabajo no habría sido posible sin la inestimable colaboración de J. A. Gualtieri, en el GSFC (NASA).

Los medios utilizados han sido cofinanciados por la Junta de Extremadura y la Comunidad Económica Europea (Proyecto TEITOSR).