

CLASIFICACIÓN DE IMÁGENES MULTIESPECTRALES MEDIANTE REDES NEURONALES CON FUNCIONES DE ACTIVACIÓN PERIÓDICAS.

C. Puig (*), E. de Miguel (**) y J.M Sopena (***)

carol.puig@upc.es

(*)*Dpto. Ingeniería del Terreno, Cartográfica y Geofísica, Universidad Politécnica de Cataluña.
C/ Gran Capità s/n. 08034 Barcelona.*

(**) *Instituto Nacional de Técnica Aeroespacial*

Carretera de Ajalvir s/n.28850 Torrejón de Ardoz. Madrid.

(***) *Laboratorio de Neurocomputación. Universidad de Barcelona
P. Vall d'Hebron, 171. 08035 Barcelona.*

Resumen: Las redes Neuronales artificiales han demostrado ser un elemento útil en la clasificación de imágenes multiespectrales especialmente cuando se desconfía de la normalidad de las cubiertas a clasificar. En este trabajo se muestra el diseño de dos redes neuronales artificiales de aprendizaje supervisado mediante el algoritmo de *backpropagation* que ha permitido la clasificación de una imagen multiespectral del sensor Daedalus Airborne Thematic Mapper (ATM), propiedad del Instituto Nacional de Técnica Aeroespacial (INTA), captada de una zona del Parque Nacional de Doñana, con funciones de activación monótonas y periódicas en cada una de ellas y que han producido resultados francamente satisfactorios.

Palabras clave: Redes neuronales, función de activación, Doñana, imágenes multiespectrales, clasificación.

INTRODUCCIÓN

Los métodos tradicionales de clasificación de imágenes espectrales utilizan técnicas de análisis estadístico multivariante y, por lo general, asumen el comportamiento normalizado de cada una de las diferentes cubiertas presentes en la imagen. Otra serie de técnicas de clasificación, usadas minoritariamente, pueden ser más adecuadas en ciertas circunstancias, por ejemplo cuando se desconfía de la normalidad de las clases, como es el uso de redes neuronales (Atkinson y Tatnall 1997).

En una clasificación supervisada con una Red Neuronal Artificial (red neuronal artificial *multilayer perceptron* de una capa oculta, con entradas, salidas, pesos, umbrales y funciones de activación monótonas o periódicas) entre las capas ocultas y lineal en la capa de salida, podremos aproximar cualquier tipo de función y discriminar regiones de forma arbitraria (contornos arbitrarios (Bishop 1995).

El proceso de clasificación de imágenes multiespectrales por Redes Neuronales Artificiales se

divide en dos partes: la primera consiste en el diseño, entrenamiento, test y validación de la red; la segunda es la clasificación propiamente dicha de toda la imagen mediante redes neuronales artificiales. Se divide en dos fases: En la primera se diseña la red, se entrena mediante un conjunto etiquetado de píxeles, y se evalúa su comportamiento mediante un conjunto de test independiente o mediante el uso de *crossvalidation*. En la segunda fase se utiliza la red ya entrenada para la clasificación de imágenes completas con la comprobación de la bondad de la misma a través de la matriz de confusión correspondiente.

En este trabajo se muestra el diseño de dos redes neuronales artificiales de aprendizaje supervisado mediante el algoritmo de propagación hacia atrás, o más conocido como *backpropagation*, una arquitectura multicapa donde la diferencia entre ambas redes ha sido exclusivamente la función de activación utilizada. Una de ellas se ha diseñado con funciones de activación monótonas mientras que la otra red, la función de activación ha sido una función periódica para la capa oculta y una función lineal en la capa de salida (Sopena et alii. 2001).

CARACTERÍSTICAS DE LAS REDES NEURONALES ARTIFICIALES.

La arquitectura usual en la clasificación de imágenes multispectrales es la *multilayer perceptron*, también conocida como *feed-forward multilayer*. Es una red multicapa, eficaz y sólida en el ajuste de pesos que produce un error mínimo, el algoritmo de aprendizaje es el *backpropagation*.

El número de capas ocultas o *hidden layers* así como el número de nodos en dichas capas se determina experimentalmente en cada ejercicio de clasificación aunque la teoría demuestra que un nodo tipo *perceptron* simple se comporta como una función booleana [4] y solamente puede representar funciones linealmente separables, de manera que en una estructura de dos capas sin capa oculta, sólo podría representar funciones linealmente separables. Con tres capas, una de ellas oculta, se pueden identificar regiones convexas abiertas o cerradas; con una estructura de cuatro capas, dos de ellas ocultas, se pueden discriminar contornos polinomiales de más o menos complejidad dependiendo del número de nodos que contengan dichas capas. Se podría continuar incrementando el número de capas ocultas y la red aprendería problemas más complicados, pero la capacidad de generalización disminuiría así como, además, aumentaría el tiempo de entrenamiento.

FUNCIONES DE ACTIVACIÓN MONÓTONAS Y PERIÓDICAS.

Cualquier función diferenciable y continua, puede utilizarse como función de activación en el algoritmo de aprendizaje *backpropagation*, aunque la selección específica depende del problema planteado de clasificación.

Los algoritmos de clasificación actúan dividiendo el espacio de variables en regiones de manera que en cada una de estas regiones se encuentre exclusivamente los elementos que pertenezcan a ellas.

En el caso de las redes neuronales cada nodo unidad contribuye dividiendo el espacio de entrada de forma específica. Los nodos Las unidades con función de activación monótona (sigmoideas; tanto la tangente hiperbólica como la logística

sinoidal) generan un hiperplano que divide el espacio de entrada en dos regiones. La combinación de varios de estos hiperplanos pueden producir clasificaciones complejas complejas.

Un nodo con una Una unidad con función de activación no monótona puede dividir el espacio en más de dos regiones. Si la función es periódica, el nombre de regiones es infinito. Se puede demostrar que una red neuronal con funciones seno y con una capa oculta es también un aproximador universal (Sopena *et alii*, 2001). Estos autores han mostrado que con este tipo de funciones se pueden obtener mejores resultados que con las funciones de tipo monótono para algunos problemas. En este trabajo exploramos si esto es también así para el caso de la clasificación de imágenes espectrales.

Normalmente, en el algoritmo de aprendizaje backpropagation, la función de activación en la capa de salida es lineal.

APLICACIONES DE LA RED NEURONAL A IMÁGENES ATM (IMAGEN ATM UTILIZADA EN LA CLASIFICACIÓN).

La imagen utilizada para la clasificación multispectral a través de redes neuronales ha sido una imagen cedida por el Instituto Nacional de Técnica Aeroespacial (INTA) del sensor Daedalus Airborne Thematic Mapper (ATM) captada en la zona del Parque Nacional de Doñana.

El sensor ATM es un escáner multispectral con la siguiente configuración espectral TMS 1.6/2.2/8.5-13µm

Canal	Longitud de onda
1	0.42-0.45
2	0.45-0.52
3	0.52-0.60
4	0.60-0.62
5	0.63-0.69
6	0.69-0.75
7	0.76-0.90
8	0.91-1.05
9	1.55-1.75
10	2.08-2.35
11	--
12	8.50-13.0

Tabla 1.- Longitudes de onda en cada canal

El canal 1, fue excluido debido a su alto ruido.

Se realizaron dos pasadas a una altura de 2800 metros y con una resolución espacial de 6 metros. La pasada P1 une el Lucio de Mari López con el Palacio de Doñana, y la pasada P2 va del palacio de Doñana a Matalascañas. La fecha de adquisición de la imagen fue el 14 de junio de 2000.

Se eligió una subescena de la pasada 1 (374 filas por 482 columnas), que contuviese elementos fuertemente diferenciados del paisaje en la zona de Doñana para la clasificación (Figura 1).

Los vectores seleccionados para el entrenamiento de cada clase (Figura 1) se eligieron en función de las texturas producidas por unos campos cosechados que ocupan de forma regular la parte inferior derecha (clase 2 a la 7) y que tiene por frontera en la parte superior izquierda de la imagen el Lucio de Mari López (clase 1).

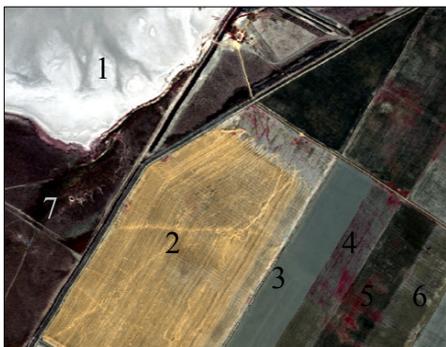


Figura 1.- Composición en falso color de la subescena utilizada para la realización de la clasificación con las clases seleccionadas para hacer la clasificación

METODOLOGÍA

Ciertos factores son importantes para la obtención de una red neuronal que pueda proporcionar resultados satisfactorios como son: los pesos iniciales, la constante de aprendizaje, la medida y la naturaleza del conjunto de entrenamiento, la arquitectura de la red y el número de iteraciones.

Las redes se han estructurado en tres capas con 40 unidades de procesado en la capa oculta. El algoritmo de aprendizaje ha sido la regla delta generalizada o *backpropagation* con una constante de aprendizaje de 0.02.

Como función de activación se utilizó en una de las redes una función sigmoidea en concreto la tangente hiperbólica y en la otra red, la función de activación fue una función periódica, la función *seno*, para la capa oculta y función lineal en la capa de salida.

El entrenamiento ha seguido las siguientes fases:

Se han seleccionado los vectores de las diferentes cubiertas, de 250 a 300 para cada clase. Los datos de entrada se han normalizado en valores comprendidos entre -1 y 1. Se inician los pesos, con valores aleatorios y pequeños. El error máximo posible de la red con la utilización de las 10 bandas espectrales, fue del 1% del máximo posible.

Una vez entrenada la red y antes de clasificar toda la imagen, se realiza un *crossvalidation* a la red y se comprueba de esta manera que la red este bien entrenada y que permita una buena generalización.

Con el fin de verificar la bondad de la clasificación de la imagen completa, se ha calculado la matriz de confusión (Chuvieco 1996) para cada imagen clasificada, obteniendo el estadístico κ y la fiabilidad global.

RESULTADOS

En la red diseñada con la función de activación *tan* y con la metodología anteriormente comentada, el resultado en el conjunto de entrenamiento fue de 98.45 % y en el conjunto de test de 98.76%. El tiempo de entrenamiento necesario para el perfecto aprendizaje de la red en 500 iteraciones fue de 22 minutos.

Una vez procesada toda la imagen se realizó la matriz de confusión (tabla 2) con otro conjunto diferente de píxeles a los utilizados en el proceso de entrenamiento, obteniendo una fiabilidad total de 93.8566% y un coeficiente kappa 0.9280.

	1	2	3	4	5	6	7	Total	Error de comisión
sin clasificar	0	0	25	10	18	6	0	59	
1	1515	0	0	0	0	0	0	1515	0.00
2	0	1406	0	0	0	0	0	1413	0.50
3	0	0	1200	0	0	0	0	1200	0.00
4	0	0	65	595	50	9	0	719	17.25
5	0	0	0	4	1507	12	1	1524	1.12
6	0	0	1	314	22	781	0	1118	30.14
7	0	0	0	0	0	0	1307	1307	0.00
Total	1515	1406	1291	923	1604	806	1308	8855	
Error de omisión	0.00	0.00	7.05	35.64	6.05	3.34	0.08		

Tabla 2.- Matriz de confusión de la imagen clasificada utilizando funciones de activación tangente hiperbólica.

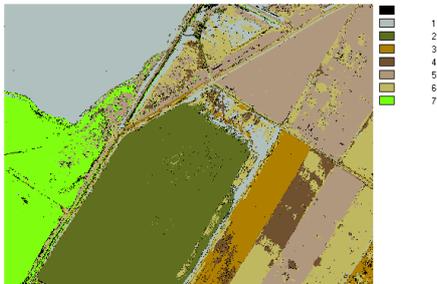


Figura 2.- Clasificación obtenida con la función de activación tangente hiperbólica.

SeEn el caso de la función seno se obtuvieron unos resultados del 100% del conjunto de entrenamiento y del 99.78 % de los píxeles de test. El tiempo utilizado en el entrenamiento fue de 15 minutos en las 500 iteraciones impuestas. La diferencia en tiempo se debe a que la derivada de la función seno es menos costosa de computar.

La matriz de confusión (tabla 3), realizada con el mismo conjunto de píxeles que en el caso de funciones de activación tangente hiperbólica, nos proporcionó unos resultados de fiabilidad total de 94.8391 % y un coeficiente kappa de 0.9394.

	1	2	3	4	5	6	7	Total	Error de comisión
sin clasificar	0	0	13	19	24	2	0	58	
1	1515	0	0	0	0	0	0	1515	0.00
2	0	1406	0	0	0	0	0	1406	0.00
3	0	0	1257	0	1	0	0	1258	1.63
4	0	0	21	595	8	12	0	626	36.62
5	0	0	0	11	1546	12	1	1570	3.62
6	0	0	0	308	25	782	0	1115	3.22
7	0	0	0	0	0	0	1307	1307	0.08
Total	1515	1406	1291	923	1604	806	1308	8855	
Error de omisión	0.00	0.00	0.00	6.55	1.53	29.81	0.00		

Tabla 3.- Matriz de confusión de la imagen clasificada utilizando funciones de activación seno.

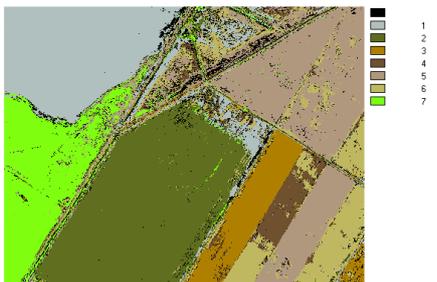


Figura 3.- Clasificación obtenida con la función de activación seno.

CONCLUSIONES

La clasificación de imágenes multiespectrales a través de redes neuronales es eficiente, comprobándose que la elección de la función de activación es importante. En otro trabajo (Puig y de Miguel 2003) hemos mostrado que su eficacia es notablemente superior a la de los métodos tradicionales.

Los resultados tal y como se intuía desde un inicio, han sido significativamente mejores en la red diseñada con funciones de transferencia periódicas y en concreto con la función sino-seno. Estas han ofrecido una convergencia más rápida y mejores resultados que las redes definidas por funciones de transferencia monótonas, en particular la función tangente hiperbólica. Las diferencias aunque pequeñas son significativas.

Como trabajo futuro tenemos previsto examinar si estas diferencias podrían aumentar en el caso de que la tarea de clasificación se hiciera más compleja al introducir más clases. También tenemos previsto examinar si la capacidad de generalización de las redes con senos es mejor en el caso de imágenes procedentes de distintos sensores.

REFERENCIAS

- Atkinson, P.M, Tatnall A.R.L. 1997. "Neural Networks in remote sensing". *International Journal of Remote Sensing*, vol. 18, 699-709.
- Bishop C.M. 1995. *Neural Networks for Pattern Recognition*. Clarendon Press. Oxford.
- Chuvieco E. 1996. *Fundamentos de Teledetección Espacial*. Rialp, Madrid.
- Puig, C y de miguel, E. 2003. *Clasificación de imágenes multiespectrales a través de redes neuronales*. 5 Setmana Geomàtica de Barcelona.
- Sopena, J.M, Romero, E y Alquézar, R. 2001. *Neural networks with periodic and monotonic activation functions: a comparative study in classification problems*. International Conference on Artificial Neural Networks. Edimburgo. Gran Bretaña.

AGRADECIMIENTOS

El presente trabajo ha disfrutado del marco proporcionado por un proyecto de la CICYT (REN 2001-1293-C02/HID). Se quiere reconocer la ayuda prestada por las personas integrantes de las siguientes instituciones Instituto Nacional de Técnica Aeroespacial y alAeroespacial, Instituto de Estudios Espaciales de Cataluña, Departamento de Lenguajes y Sistemas Informáticos (UPC) y Departamento de Ingeniería del Terreno, Cartográfica y Geofísica (UPC).