

CLASIFICACION DE POTENCIALES DE ACCION CON UNA RED NEURONAL QUE EMPLEA RETROPROPAGACION

Quizá Tomich J.

Espinosa Espinosa I.*

Maestría Ing. Biomédica. U.A.M. - Iztapalapa

*Lab. Cibernética - Fac. Ciencias. U.N.A.M.

R E S U M E N -----

Se describe el uso de una red neuronal, entrenada con el algoritmo de retropropagación, para clasificar impulsos nerviosos de tal forma que su arquitectura le permite identificarlos aun cuando se encuentren desplazados en el tiempo.

INTRODUCCION

La aparición de electrodos de metal que permiten registrar simultáneamente la actividad de un grupo de neuronas ha hecho patente la necesidad de desarrollar técnicas que faciliten la separación de las diferentes formas de onda correspondientes a los distintos impulsos. Al efectuar registros con este tipo de electrodos, diversos factores causan que la actividad de cada neurona individual presente una característica particular, por lo cual es esencial distinguir las señales de cada una de ellas para comprender mejor la conectividad funcional del sistema estudiado. Un panorama del tipo de problemas encontrados al hacer registros múltiples simultáneos con diferentes configuraciones de electrodos puede encontrarse en Gerstein et al. (1983).

Una manera de clasificar estas señales es por la forma de la espiga del impulso nervioso, pero es necesario tener en cuenta que los registros siempre estarán contaminados con ruido, bien sea este proveniente de fuentes externas o de las interacciones de la misma señal registrada con señales provenientes de neuronas distantes.

Generalmente, las técnicas de clasificación empleadas están basadas en la teoría de los filtros análogos y digitales así como en el reconocimiento de patrones estadístico (Wheeler y Heetderks, 1982; Yang y Shamma, 1988). Los métodos más sofisticados obtienen mejores resultados, aunque muchas veces su complejidad no permite que sean utilizados en tiempo real. Con el fin de encontrar una alternativa confluente y de posible aplicación en tiempo real, se decidió aprovechar la habilidad de las redes neuronales artificiales en el reconocimiento de patrones adaptivo y emplear una de ellas en la tarea de clasificación.

REDES NEURONALES ARTIFICIALES

Las redes neuronales artificiales están inspiradas en los sistemas biológicos en los que un gran número de células nerviosas, que individualmente funcionan de una forma más bien torpe, al actuar colectivamente pueden efectuar tareas extremadamente complejas, que muchas veces no han podido ni siquiera ser imitadas por las computadoras existentes más avanzadas. Una red neuronal artificial puede definirse como un sistema computacional compuesto por elementos procesadores simples con un alto grado de conectividad entre ellos, que procesa información mediante su respuesta dinámica a señales externas (Hecht-Nielsen, 1987). Actualmente, puede encontrarse una abundante literatura sobre ellas (Grossberg, 1988; Kohonen, 1989; Rumelhart y McClelland, 1986, etc.).

Los elementos procesadores (fig. 1), que son los equivalentes a las neuronas biológicas, reciben señales de entrada ya sea del medio exterior, de otros elementos o incluso de sí mismos. Estas señales de entrada son ponderadas por unas variables conocidas como pesos y causan una respuesta del elemento. Tal respuesta está determinada por la función de transferencia característica del elemento; dichas funciones pueden tomar valores discretos (por ejemplo, 0 ó 1), como en las de umbrales lógicos, o valores continuos si se emplea una función lineal o, lo que es más frecuente, una de tipo sigmoidal, como la función logística o la tangente hiperbólica. Los elementos se conocen con el nombre de nodos, neuronos o neuronas. Continuando con la

analogía biológica, la respuesta del elemento es equivalente a la frecuencia de disparo, la sumatoria de las señales de entrada ponderadas al voltaje de soma y los pesos a la intensidad de la conexión sináptica. Los pesos pueden ser negativos, positivos o cero, dependiendo de si la sinapsis es inhibitoria, excitatoria o indiferente, respectivamente.

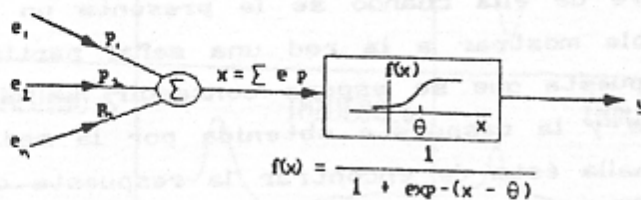


Fig. 1. Esquema de un elemento procesador. Las e_i representan a la señal de entrada. Los pesos corresponden a los p_i . La función de transferencia mostrada es la ecuación logística.

La salida de cada uno de los elementos procesadores va a otros elementos o al medio exterior. La interconectividad de los elementos de la red determina su arquitectura.

Una de las características distintivas de estos sistemas es su habilidad para aprender, por lo cual pueden emplearse en el reconocimiento de patrones. Sucintamente, la red es enfrentada a una serie de patrones de entrada de diferentes clases y debe extraer de ellos la información relevante para, posteriormente, poder clasificar de forma correcta patrones nuevos, en otras palabras, una red neuronal útil en el reconocimiento de patrones debe ser capaz de manifestar un comportamiento cognoscitivo que pueda definirse como "generalización" (Anderson, 1988).

MÉTODOS

El algoritmo de retropropagación (Rumelhart *et al.*, 1986) es uno de los métodos de aprendizaje más populares. Aunque puede aplicarse a redes con arquitecturas diferentes, lo usual es considerar una red con la topología que se describe a continuación. Los elementos procesadores se hayan dispuestos en diferente capas; una de entrada, encargada de recibir la señal del medio externo y distribuirla a la

capa siguiente, una capa de salida que comunica al exterior la respuesta de la red a la señal de entrada y una o más capas intermedias. El flujo de información es unidireccional, de la capa de entrada a la de salida. Cada capa está conectada únicamente con las capas inferior y superior inmediatas, y no hay conexiones entre los elementos de una misma capa. La arquitectura descrita es conocida como de capas múltiples con propagación progresiva.

Esta estructura particular de relación con el ambiente externo permite la intervención de un "instructor", que puede indicar a la red la respuesta que se quiere de ella cuando se le presenta un estímulo determinado. Así, es posible mostrar a la red una señal particular de entrada junto con la respuesta que se espera conseguir. La diferencia entre este valor esperado y la respuesta obtenida por la red es una medida de cuán lejos se halla ésta de encontrar la respuesta correcta. Si se define esta diferencia como una función de error, el aprendizaje consiste en minimizar el error. El algoritmo de retropropagación busca la reducción gradual del error cuadrático medio, considerándolo como una función de los pesos de la red. Matemáticamente, es un método de gradiente descendente. El procedimiento implica dos pasos a través de la red. En el primero, una señal (patrón) es presentada a la capa de entrada y, usando los valores corrientes de los pesos, se calcula la respuesta de los elementos de las diferentes capas, empezando por aquellas más cercanas a la entrada y avanzando hasta llegar a la de salida. Una vez obtenida la respuesta de la capa de salida, se compara ésta con la respuesta esperada y se tiene el error. La derivada del error es propagada ahora hacia la entrada, durante el segundo paso, y el valor de todos los pesos es ajustado con el fin de reducir el error. Esto se repite muchas veces con todas las señales que se estén empleando en el entrenamiento, hasta que la red obtiene las respuestas deseadas para las diferentes clases.

En la figura 2 aparecen ocho clases diferentes de potenciales de acción que se desea clasificar con la red. Inicialmente, se construye un conjunto de entrenamiento, que está formado por diferentes ejemplares de cada una de las clases, así como de la respuesta esperada para cada uno de ellos. Al comenzar el entrenamiento, los pesos son inicializados con valores aleatorios pequeños. Una vez

terminada la fase de aprendizaje, la eficacia de la red como clasificadora se verifica con un conjunto de prueba, que está compuesto por patrones que no han sido presentados a la red durante el periodo de entrenamiento.

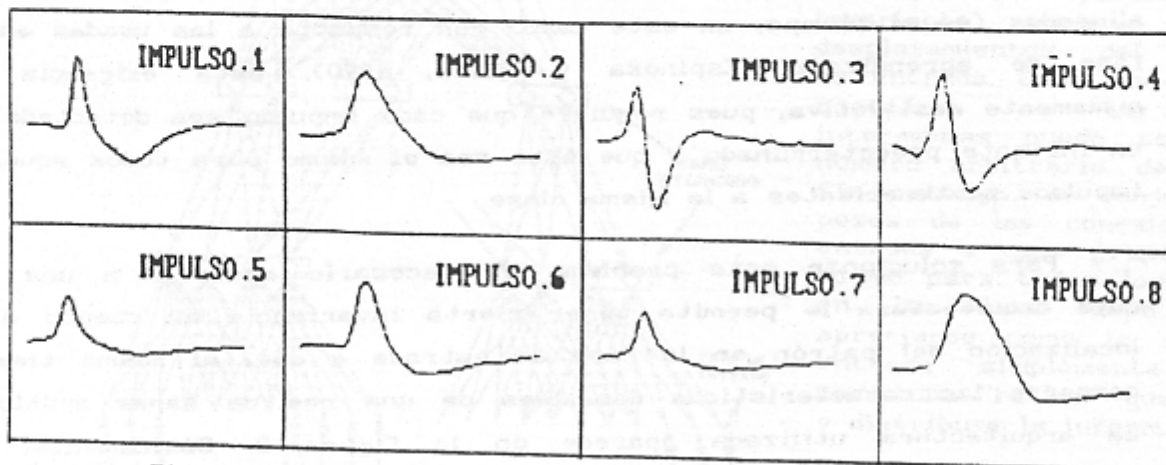


Fig. 2. Ejemplos de las ocho clases diferentes de impulsos a ser clasificadas.

Los patrones de entrenamiento usados en este caso son señales formadas por 128 puntos, que corresponden a las muestras obtenidas por un convertidor analógico-digital durante la operación de registro. Se decidió trabajar con los datos sin someterlos a ningún procesamiento previo, a excepción de un escalamiento con el objeto de limitar sus valores dentro de cierto rango, por ejemplo en el intervalo $(-1, 1)$. En consecuencia, la capa de entrada tiene 128 elementos.

La función de transferencia de los elementos de las capas de salida e intermedias es la función logística, mostrada en la figura 1, por lo que sus salidas varían entre cero y uno. En la capa de salida se usó un elemento por cada clase, de tal forma que sólo uno de ellos se activa a la vez cuando se ejecuta una clasificación. Se considera que un nodo está activo si su salida es mayor que 0.5.

El número de capas intermedias y el de elementos en ellas fue variable, aunque la mayor parte de las veces se trabajó con dos capas intermedias, puesto que, aparentemente, no es preciso usar más para generar las hipersuperficies e hipervolumenes de decisión necesarios para lograr la separación de las diferentes clases (Lippmann, 1987).

Al comienzo se empleó una configuración convencional, donde los elementos de una capa están interconectados con todos y cada uno de los elementos de las capas adyacentes. Al efectuar la evaluación de la red con patrones de prueba, se encontró que su mayor defecto es la incapacidad de reconocer patrones que se encuentren desplazados ligeramente en la capa de entrada, lo cual implica que para obtener un buen grado de reconocimiento, las señales deben estar perfectamente alineadas (en el tiempo, en este caso) con respecto a las usadas en la fase de aprendizaje (Espinosa y Quiza, 1990). Esta exigencia es sumamente restrictiva, pues requiere que cada impulso sea detectado en un instante predeterminado y que éste sea el mismo para todos aquellos impulsos pertenecientes a la misma clase.

Para solucionar este problema es necesario recurrir a una red cuya arquitectura le permita tener cierta invariancia en cuanto a la localización del patrón en la capa de entrada y que, al mismo tiempo, conserve las características deseables de una red de capas múltiples. La arquitectura utilizada aparece en la figura 3. Básicamente, las capas intermedias se han organizado en grupos de nodos que reciben sus entradas sólo de una región de la capa inferior, con un traslape entre las regiones abarcadas por cada grupo. La última capa intermedia está conectada totalmente con la de salida. Esta disposición puede interpretarse de dos maneras, ya sea como si se aplicase un retardo temporal diferente a cada una de las señales de entrada de los nodos de un grupo (Waibel et al., 1989) o como si cada grupo de nodos fuese un detector de características (Bottou et al., 1990). Con el fin de conseguir la invariancia con respecto al desplazamiento del patrón de entrada, debe asegurarse que cada grupo de nodos identifique el mismo tipo de secuencias temporales o características morfológicas. Esto significa que los grupos dentro de una capa intermedia no son sino duplicados que se han desplazado, temporal o espacialmente, y, por lo tanto, comparten el mismo conjunto de pesos.

Se utilizó un conjunto de entrenamiento con 240 patrones, treinta de cada clase. Se agregó ruido gaussiano, punto a punto, a los diferentes ejemplos de cada clase. Como conjunto de prueba se empleó uno con 160 patrones, distintos a los de entrenamiento.

Las sesiones de aprendizaje se hicieron en una computadora

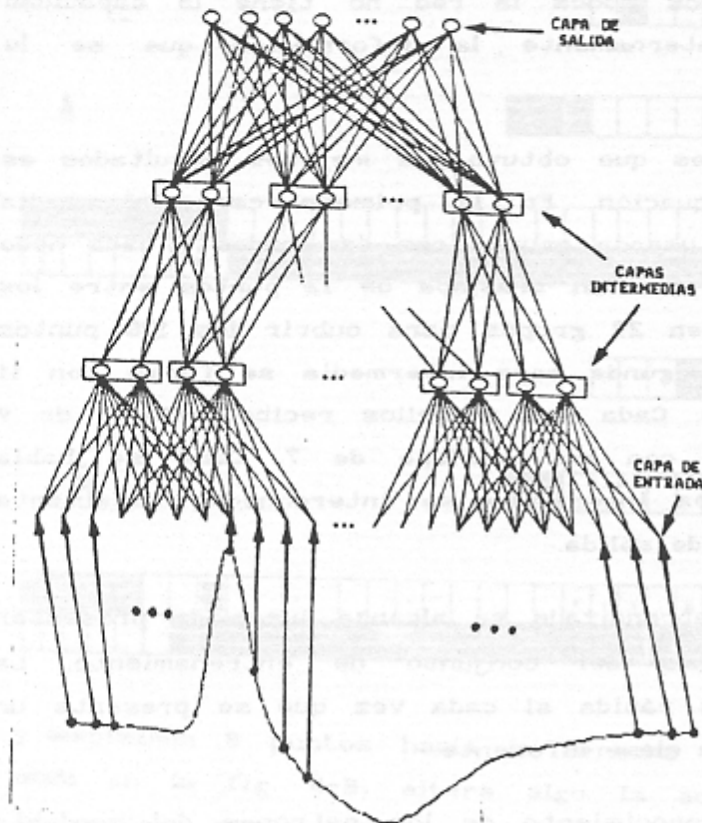


Fig. 3. Diagrama esquemático de la arquitectura de una red que es (parcialmente) invariante respecto a los desplazamientos del patrón de entrada. Cada uno de los recuadros en las capas intermedias puede tener un número arbitrario de nodos. Sin embargo, el conjunto de pesos de las conexiones de entrada a cada grupo es el mismo para todos los grupos en la misma capa. Puede apreciarse como la capa de entrada simplemente sirve como una "retina" que recibe y distribuye la información.

Stardent Titan, con dos procesadores que alcanzan que alcanzan los 16 MFLOPS, cada uno. Para las ocho categorías diferentes, una sesión típica toma de unos 4 a 5 minutos de tiempo de CPU, lo cual significa unas 2 ó 3 horas para el usuario, dependiendo de la carga del sistema. El tiempo de entrenamiento aumenta si se desean reconocer más categorías y fácilmente puede llegar a ser de varios días.

Con el fin de mantener igual el juego de pesos compartidos por los grupos de nodos de las capas intermedias, las variaciones calculadas para cada uno de los pesos en los diferentes grupos se promedian y el resultado se utiliza para actualizar sus valores.

RESULTADOS

Variando la cantidad de nodos de los grupos de las capas intermedias al igual que el número de conexiones de entrada y el traslape entre los grupos, se estudió la efectividad de la red como clasificadora de patrones, así como la velocidad de aprendizaje. En general, cuantos

más nodos se empleen en las capas intermedias, más rápido será el aprendizaje. Sin embargo, si se utilizan demasiados nodos, la red memoriza el conjunto de entrenamiento y no descubre las características propias de cada clase, por lo que al ser expuesta a los patrones de prueba su rendimiento es bastante pobre. Por otro lado, si se emplean muy pocos nodos la red no tiene la capacidad suficiente para codificar internamente la información que se le presenta.

Una de las configuraciones que obtuvo los mejores resultados es la que se delinea a continuación. En la primera capa intermedia después de la de entrada, se usaron grupos con dos nodos y cada nodo recibe 16 señales de entrada. Con un traslape de 12 puntos entre los diferentes grupos, se requieren 29 grupos para cubrir los 128 puntos de la señal de entrada. La segunda capa intermedia se formó con 11 grupos, de un nodo cada uno. Cada uno de ellos recibe señales de 9 grupos de la capa inferior, con un traslape de 7. Como se había indicado antes, la última capa intermedia se interconecta totalmente con los ocho nodos de la capa de salida.

En la red descrita, el aprendizaje se alcanza luego de presentar unas 850 veces los patrones del conjunto de entrenamiento. La disminución del error es más rápida si cada vez que se presenta un patrón éste corresponde a una clase diferente.

La eficiencia en el reconocimiento de los patrones del conjunto de prueba es un poco mayor que el 95%. Se encontró que para mejorar la respuesta de la red a patrones desplazados, ejemplos de éstos deben ser incluidos en el conjunto de entrenamiento. De hecho, la mayor parte del tiempo de entrenamiento es debida a la presencia de los patrones desplazados, mientras que el ruido es rápidamente identificado como no esencial para la caracterización de las clases.

La extracción de las características distintivas de los patrones se lleva a cabo por las capas intermedias y el proceso de ajuste de los pesos representa una especie de auto-organización en la actividad de sus elementos. Sin embargo, la mayoría de las veces es difícil interpretar lo que los estados internos codifican. En las figuras 4 y 5 se muestra la actividad de las capas intermedias y de salida para diferentes patrones de entrada. Puede apreciarse como la señal ruidosa

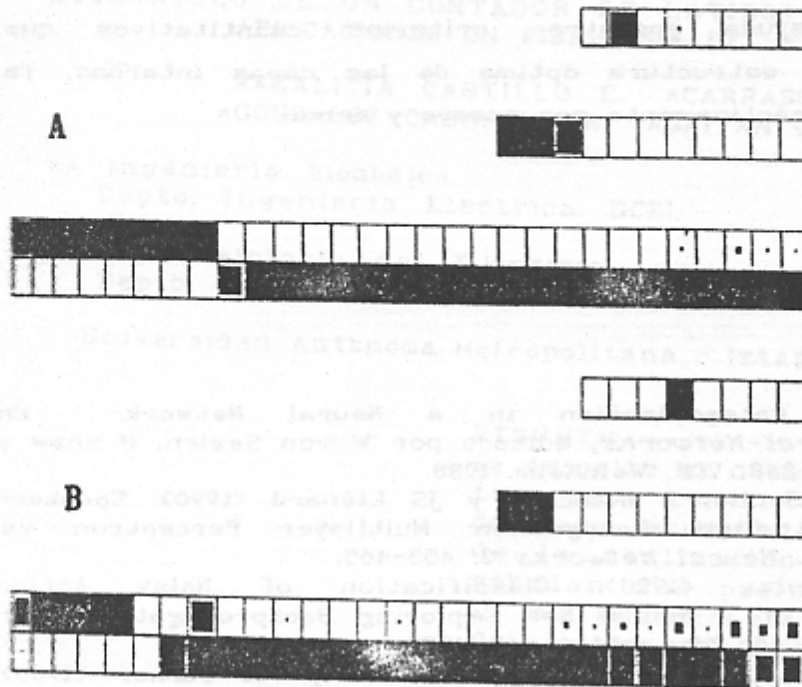


Fig. 4. Estados de las capas de salida e intermedias cuando se presentan los patrones de la clase 7 (A) y 5 (B) que se muestran en la fig. 2. Cada cuadro es un nodo y el sombreado indica el estado de su salida, menor que 0.1 en blanco y mayor que 0.9 en negro. Los grupos de nodos en la primera capa intermedia se han dibujado de forma vertical. La capa de salida aparece en la parte superior.

y desplazada 8 puntos hacia la derecha (Fig. 5), con respecto a la usada en la fig. 4-B, altera algo la activación de la primera capa intermedia, pero la actividad de las otras capas permanece prácticamente igual.

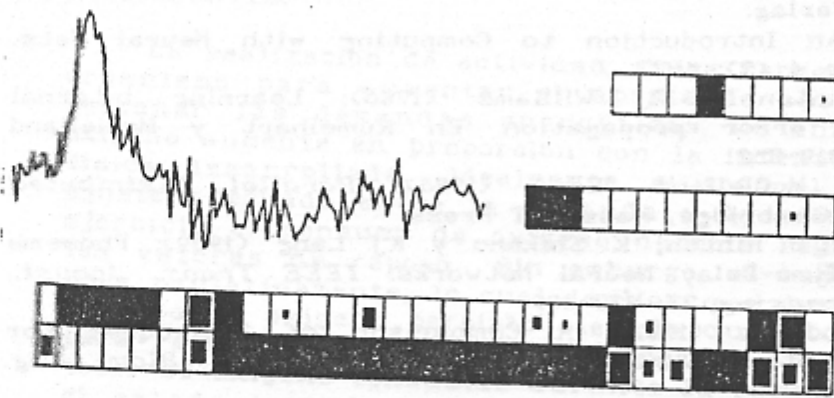


Fig 5. Estados de las capas de salida e intermedias cuando se presenta una señal ruidosa y desplazada de la clase 5.

CONCLUSIONES

Resulta evidente que la red estudiada puede ser útil en el reconocimiento de patrones en tiempo real, puesto que puede manejar casi directamente la información obtenida en el proceso de adquisición

de datos. Aunque el entrenamiento puede requerir el empleo de una computadora potente, una vez concluida esa etapa, la labor de reconocimiento se puede efectuar en una computadora menos sofisticada, incluso en una computadora personal.

Sería de gran ayuda encontrar criterios cuantitativos que permitan seleccionar la estructura óptima de las capas interñas, ya que por ahora se hace prácticamente por ensayo y error.

BIBLIOGRAFIA

- Anderson, JA (1988): Categorization in a Neural Network. En *Organization of Neural Networks*, editado por W von Seelen, G Shaw y UM Leinhos. Págs. 319-343. VCH. Weinheim. 1988.
- Bottou, L, F Fogelman Soulié, P Blanchet y JS Liénard (1990): Speaker-Independent Isolated Digit Recognition: Multilayer Perceptron vs. Dynamic Time Warping. *Neural Networks* 3: 453-465.
- Espinosa, I y J Quiza (1990): Classification of Noisy Action Potentials by Means of a Neural Net Employing Backpropagation. *Soc. Neurosci. Abstr.*, vol. 16, 2da. parte, pág. 1092.
- Gerstein, GL, MJ Bloom, I Espinosa, S Evanczuck y MR Turner (1983): Design of a Laboratory for Multineuron Studies. *IEEE Trans. Sys., Man, Cyb.* 13: 668-675.
- Grossberg, S (Ed.) (1988): *Neural Networks and Natural Intelligence*. Cambridge, Mass. MIT Press.
- Hecht-Nielsen, R (1987): Neurocomputer applications. En *Neural Computers*, editado por R Eckmiller y G von der Malsburg. Heidelberg. Springer-Verlag. 1987.
- Kohonen, T (1989): *Self-Organization and Associative Memory*. 3a. ed. New York. Springer-Verlag.
- Lippmann, RP (1987): An Introduction to Computing with Neural Nets. *IEEE ASSP Magazine* 4 (2): 4-22.
- Rumelhart, DE, GE Hinton y RJ Williams (1986): Learning internal representations by error propagation. En Rumelhart y McClelland (1986). Vol. 1, págs. 318-362.
- Rumelhart, DE y JL McClelland (Eds.) (1986): *Parallel Distributed Processing*. 2 vols. Cambridge, Mass. MIT Press.
- Waibel, A, T Hanazawa, G Hinton, K Shikano y KJ Lang (1989): Phoneme Recognition Using Time-Delay Neural Networks. *IEEE Trans. Acoust., Speech, Signal Processing* 37: 328-339.
- Wheeler, B y W Heetderks (1982): A Comparison of Techniques for Classification of Multiple Neural Signals. *IEEE Trans. Biom. Eng.* 29: 752-759.
- Yang, X y S Shamma (1988): A Totally Automated System for the Detection and Classification of Neural Spikes. *IEEE Trans. Biom. Eng.* 35: 806-816.