

# Reconocimiento y clasificación de fibras optimizando una red neuronal basada en el modelo neocortical

## Fiber recognition and classification optimizing a neural network based on the neocortical model

Presentación: 00/00/0000

Doctorando:

**Marcelo Arcidiácono**

Universidad Tecnológica Nacional – Facultad Regional Córdoba  
marceloarcidiacono@gmail.com

Director/es:

**Eduardo Destefanis**

### Resumen

En el contexto de la sustentabilidad productiva de fibras textiles, la oportunidad de contar con un método accesible y económico para reconocer y clasificar tipos específicos y obtener así una medida de calidad confiable por parte de productores rurales, es de vital importancia. En el presente trabajo se optimiza un método de reconocimiento y clasificación de objetos basado en una Memoria Temporal Jerárquica para lograr dicho objetivo. La Memoria Temporal Jerárquica, inspirada en la teoría de la memoria predicción del cerebro humano, consiste en una estructura de árbol de nodos computacionalmente conectados los cuales utilizan un conjunto particular de reglas para memorizar objetos que aparecen en diferentes orientaciones [Hawkins & Ahmad]. En concordancia con este modelo, se aplican algoritmos basados en los mecanismos de visión humana para preprocesar las imágenes de entrada. Así, las imágenes de entrada son sometidas a un proceso basado en un modelo matemático para resaltar las características visuales más relevantes (tal como lo hace el cerebro humano). En comparación con el método original de la Memoria Temporal Jerárquica, los resultados experimentales obtenidos exhiben una mejora en el rendimiento y precisión, logrando además un entrenamiento más rápido que conserva la robustez original.

Palabras claves: fibra, redes neuronales.

### Abstract

In the productive sustainability of animal textil fiber framework, having a quick, safe and inexpensive method that allows to obtain a measure of fiber quality to compete in local and international markets, is highly beneficial for farmers. In this work, a method of recognition and classification of objects based on a Hierarchical Temporary Memory is optimized for that purpose. The Hierarchical Temporary Memory, inspired by the theory of prediction memory of the human brain, consists of a tree structure of computationally connected nodes which have a particular set of rules for memorizing objects that appear in different orientations [Hawkins & Ahmad]. The input images were subjected to a preprocessing based on a mathematical model to highlight the most relevant visual characteristics. The experimental results demonstrated an improvement in performance and precision.

Keywords: fiber, neural network.

## Introducción

Tanto la identificación como la caracterización de pelos y fibras de origen animal, adquiere importancia relevante en el análisis y estudio de sus propiedades para usos textiles e industriales, incluso artesanales [Ford]. Se estima que más de un millón de pequeños productores de los Andes centrales en Sudamérica tienen alpacas y llamas como principal medio de subsistencia [Mueller]. Estos animales proveen carne, leche, fibra, energía de transporte y guano y, además, constituyen un elemento importante de la identidad cultural de sus pueblos. Poblaciones definidas de estos camélidos califican para ser capturadas, esquiladas y liberadas generando un ingreso adicional a las comunidades en que viven. El aumento de la producción de fibras y demás productos de camélidos sudamericanos, a la vez de preservar un recurso genético animal crítico y los valores culturales asociados y mejorar la calidad de vida de muchos pequeños productores, debe ser parte de una estrategia global de inversión sostenida en investigación y desarrollo. En la República Argentina, el Programa Nacional Fibras Animales [Mueller] considera valiosa la comercialización, producción, e industrialización de lana, mohair, cashmere, llama, guanaco y vicuña. El valor de la fibra textil está dado, fundamentalmente, por su finura promedio además de otras propiedades que hacen a establecer su cotización tales como el índice de confort PF (Prickle Factor) que constituye el porcentaje de fibras con diámetros mayores a 32 micrones, la presencia o ausencia de medulación<sup>1</sup>, el crimpado<sup>2</sup> y la forma y altura de las escamas [Adot]. Para determinar una medida satisfactoria de calidad de la fibra textil de origen animal, además de tener en cuenta defectos obvios como la pigmentación y la presencia de fibras atípicas o meduladas [Cancio], la característica de mayor importancia es el diámetro medio. Fibras más finas tienen más aplicaciones industriales y en consecuencia tienen mayor valor económico [Mueller]. Reconocer y clasificar apropiadamente cada tipo de fibra, para medir sus características biomecánicas más relevantes, constituye el objetivo de este trabajo.

### Estrategia basada en el modelo del neocórtex

El problema del reconocimiento y clasificación de objetos en imágenes parece simple a primera vista, pero a menudo es difícil de resolver. No hay una manera obvia de manejar estos problemas y existen muchas técnicas diferentes. El reconocimiento visual de patrones, tal como la lectura de caracteres o la discriminación de formas, es una tarea sencilla y natural en seres humanos. La verdadera dificultad consiste en diseñar un algoritmo computacional capaz de hacer esto de la misma manera. Quizás, una de las mejores estrategias sea aprender del propio cerebro, estudiando el mecanismo con el que se procesa la información visual y tratando de usarlo en el principio de diseño de nuevos algoritmos. Más precisamente, cómo puede sintetizarse una red neuronal que tenga la misma habilidad que el cerebro humano. Un resultado de dicha aproximación en el reconocimiento de patrones es el llamado «neocognitrón» [Fukushima].

En el área visual del cerebro, se encontraron neuronas que responden selectivamente a características locales de un patrón visual tales como líneas y bordes y, en particular, orientaciones [Huble]. En la zona más alta de la corteza visual, existen células que responden selectivamente a ciertas figuras como círculos, triángulos, cuadrados o incluso al rostro humano [Bruce]. De acuerdo a esto, el sistema visual parece tener una estructura jerárquica en la que primero se extraen características simples de un patrón estímulo y luego se integran características más complejas. En esta jerarquía, una celda en una capa superior generalmente recibe señales desde un área más amplia de la retina siendo insensible a la posición del estímulo.

En este trabajo proponemos un método mejorado de reconocimiento y clasificación de objetos basado en la técnica de Memoria Temporal Jerárquica (Hierarchical Temporal Memory HTM) inspirada en la teoría de predicción del cerebro humano [Hawkins & Ahmad]. La Red HTM fue propuesta por Jeff Hawkins y modela algunas de las estructuras y propiedades algorítmicas del neocórtex. La Red HTM es un modelo biométrico basado en la teoría memoria-predicción de la función cerebral descrita en [Hawkins & Blakeslee]. Se trata de un método para descubrir e inferir las causas de alto nivel de los patrones y secuencias de entrada observados, construyendo así un modelo cada vez más complejo de la realidad. Una Red HTM combina y se extiende a enfoques utilizados en redes bayesianas y algoritmos de agrupación espacial y temporal mediante el uso de una jerarquía en forma de árbol. La Red HTM se ha utilizado con éxito para el reconocimiento de objetos [Bundzel] [George] ya que el proceso de construcción de mapas de la realidad tiene características intrínsecamente similares a la de identificación de objetos. Al tratar el problema de identificación de fibras como un problema de reconocimiento espacial, proponemos resolverlo con una Red HTM, usando como base los mecanismos de la visión humana en la implementación de un algoritmo que permita pre procesar las imágenes de entrada. De esta forma, las imágenes de entrada se someten a un procesamiento basado en un modelo matemático que permite realzar las características visuales más sobresalientes, aquellas que la vista humana mayormente detecta.

### Estructura de un HTM

La Memoria Temporal Jerárquica está inspirada en las redes bayesianas, sin embargo, su principal característica es la de modelar la estructura algorítmica del neocórtex [Mittal]. Una Red HTM es un sistema organizado jerárquicamente, dividido en varios niveles compuestos de nodos computacionales adjuntos. Al ascender, el número de nodos disminuye siendo tan solo uno en el nivel superior. El número de nodos en el nivel  $k$  es  $2^{2n-k}$ , donde  $n$  es el número total de niveles en la red. El nivel más bajo corresponde a las imágenes que se presentan a la Red. Las imágenes de entrada se dividen a su vez en cuadros de  $m \times m$  píxeles haciendo corresponder cada cuadro de la imagen de entrada a un nodo del primer nivel, así, los nodos del primer nivel reciben entradas de áreas específicas llamadas campos

receptivos. El flujo de inferencia de uno a otro nodo, sigue una trayectoria ascendente. Los nodos en niveles inferiores pueden encontrar causas pertenecientes a una escala de tiempo y entrada limitada, mientras que los nodos en niveles superiores encuentran causas que pertenecen a una escala de tiempo y entrada más amplia. Independientemente del nivel al que pertenecen, cada nodo sigue el procedimiento algorítmico descrito en [George & Jaros] dividido en dos módulos diferentes: el módulo espacial y el módulo temporal y realiza dos operaciones distintas: la primera, entrenamiento, incluye procesamiento temporal y espacial; la segunda, inferencia, produce las salidas que alimentan los nodos en los niveles superiores.

### **El módulo espacial**

Un nodo en el primer nivel recibe un cuadro de tamaño  $m \times m$  píxeles del correspondiente campo receptivo. Este cuadro se mapea en un vector de entrada  $\mathbf{q} = (q_1, q_2, \dots, q_{m \times m})$  donde los  $q_i$  corresponden al valor de cada pixel. El nodo es expuesto secuencialmente a las imágenes de entrenamiento y aprende el espacio de cuantificación del vector de entrada. El subespacio de dichos vectores, eventualmente añadidos al nodo, constituye los centros de cuantificación. Este procedimiento de aprendizaje se implementa mediante un algoritmo simple, gobernado por un parámetro de umbral  $D$ , que corresponde a la distancia euclidiana mínima por encima de la cual un vector de entrada se considera diferente de los centros de cuantificación existentes. Cada nuevo vector de entrada se comprueba contra el centro de cuantificación con una distancia menor a  $D$ , en ausencia de tal centro, el nuevo vector se agrega al nodo convirtiéndose en un nuevo centro de cuantificación, de lo contrario, se ignora. El valor de umbral de  $D$  debe elegirse con cuidado, ya que un valor demasiado pequeño puede conducir a un nodo con demasiados centros de cuantificación. Por otro lado, un valor demasiado grande puede conducir a un nodo con centros de cuantificación correspondientes a vectores de entrada muy diferentes entre sí. Este procedimiento de aprendizaje converge cuando la tasa de incorporación de nuevos centros de cuantificación es mucho menor a un valor predefinido. Una vez que el procedimiento de aprendizaje concluye, el módulo espacial está entrenado y puede producir salidas que alimenten el módulo temporal.

### **El módulo temporal**

El módulo temporal aprende a partir de las salidas del módulo espacial. Por cada nuevo vector de entrada presentado a un nodo se calculan las distancias euclidianas entre dicho vector y los centros de cuantificación. La probabilidad de que el vector de entrada pertenezca a cada centro alimenta el módulo temporal para formar la llamada Matriz de Adyacencia Temporal (MAT). Si el módulo espacial consta de  $N$  centros de cuantificación, se crea primero una matriz  $T \ N \times N$  en cero. Las filas y las columnas de la matriz corresponden a los centros activados en los tiempos  $t$  y  $t + 1$ , respectivamente. De esta forma, el elemento  $T_{ij}$  de la MAT se incrementa en uno cuando el centro de cuantificación correspondiente es activado por un vector de entrada en los tiempos  $t$  y  $t + 1$  y la operación se repite hasta que la matriz logra suficiente estabilidad. El significado físico de este procedimiento es que los centros de cuantificación con tiempos próximos se agrupan juntos en  $T$ . Luego la matriz se usa para aprender una combinación de cadenas de Markov de primer orden, definidas sobre el set de centros de cuantificación.

### **Agrupamiento temporal**

El siguiente paso incluye la partición de la MAT, que es una matriz grande de transición. El objetivo principal de esta operación es agrupar los centros de cuantificación en subgrupos temporales coherentes. Cada subgrupo corresponde a una cadena de Markov e incluye centros de cuantificación que es probable ocurran secuencialmente en el tiempo [George & Hawkins]. Si consideramos a la MAT como una red dirigida ponderada, este problema se trata como una partición gráfica y el algoritmo que ha sido adoptado como solución es el descrito por George y Jaros en [George & Jaros]:

1. Localizar el centro de cuantificación con la mayor conectividad.
2. Seleccionar los  $M$  centros de cuantificación con mayor conectividad con el centro del paso anterior.
3. Repetir el paso anterior para cada nuevo centro de cuantificación agregado.
4. Crear un nuevo grupo compuesto por los centros que surgen de los pasos 2 y 3.
5. Repetir, desde el paso 1, hasta que todos los centros de cuantificación hayan sido agregados a un grupo.

La salida del módulo temporal será un vector que contendrá el número de grupos temporales creados para un nodo. El entrenamiento termina al completar la partición gráfica y el nodo es capaz de hacer inferencia.

En este trabajo, tanto el módulo espacial como el módulo temporal se reemplazan con rutinas optimizadas que se proponen a continuación. En el módulo espacial el algoritmo propuesto permitirá almacenar nuevos centros de cuantificación de manera más eficiente. En cuanto al módulo temporal, la matriz  $T$  se reemplaza por una matriz de autocorrelación que permite potenciar el análisis factorial.

### **Descripción del algoritmo propuesto**

La contribución de este trabajo es doble. Por un lado, sobre las imágenes de entrada, cortes longitudinales de fibra textil, se aplica un algoritmo de mejoramiento cuyo objetivo es perfeccionar topológicamente la estructura de la

imagen. Durante este proceso, se extraen las coordenadas de la imagen que corresponden a la información de las regiones relevantes. Por otro lado, se utiliza sólo la información extraída para entrenar una Red HTM.

### **Filtro de Gabor modificado**

El método de mejoramiento de la imagen de entrada adoptado en este trabajo se basa en un filtro de Gabor, técnica que combina dos capacidades de filtrado: una en el dominio espacial y otra en el dominio de frecuencia. Es esta característica específica la que se explota para mejorar la estructura de la imagen. La base de esta técnica se encuentra en una de las aplicaciones más recientes en las que se filtra con Gabor en el dominio de frecuencia para mejorar la estructura de huellas dactilares [Li]. Tradicionalmente se asume que la estructura paralela de crestas y valles, en huellas dactilares, recorrida ortogonalmente a la orientación local de las líneas de crestas, conforman una onda plana de forma sinusoidal combinada con ruido. Esto supone que la señal ortogonal a la orientación local de las crestas, es aproximadamente una onda sinusoidal discreta, con lo que sería suficiente sintonizar el filtro a la orientación local y a la frecuencia de la onda conformada por las crestas y valles para poder eliminar el ruido y restablecer la estructura paralela de crestas y valles de la huella original. Siguiendo el trabajo de Yang et al. en [Yang] se aplica este enfoque, teniendo en cuenta que las fibras pueden ser consideradas como dos, tres y hasta cuatro crestas paralelas según se trate de fibras meduladas o no meduladas (en algunos casos, dependiendo del ancho de la médula, ésta se presenta en la imagen agregando una o dos crestas paralelas a los bordes de la fibra).

### **Módulo espacial modificado**

El módulo espacial es expuesto secuencialmente a las imágenes de entrada modificadas. El objetivo es que el módulo espacial aprenda un subconjunto representativo de los vectores de entrada. Los vectores de entrada almacenados son los centros de cuantización que codifican el conocimiento de la red. Es muy importante que estos centros se seleccionen cuidadosamente para garantizar que el módulo espacial aprenda un conjunto finito de centros de cuantización a partir de un número virtualmente infinito de vectores de entrada. Debe alcanzarse en este punto un equilibrio entre el número de centros de cuantización almacenados y el resto de espacio. El algoritmo propuesto permite almacenar nuevos centros de cuantificación de manera más eficiente.

### **Módulo temporal modificado**

Lo expresado anteriormente respecto al módulo temporal es aplicable en aplicaciones donde la entrada corresponde a una secuencia basada en el tiempo tales como imágenes de video, etc. Los vectores de entrada que se presentan en el campo receptivo de un nodo a lo largo del tiempo corresponden al mismo contenido desplazado por una pequeña porción de píxeles. En tales casos los vectores de entrada con proximidad temporal también tendrían espacio. No obstante, en el reconocimiento y problemas de categorización, donde el sistema debe entrenarse a partir de un conjunto de imágenes sin proximidad temporal, la aproximación no tiene un significado físico. Es raro que las imágenes de entrada presentadas a un nivel más bajo, de forma sucesiva, compartan información común. Como consecuencia de esto, el campo receptivo de un nodo recibe vectores que pueden diferir notablemente. Como sea, la matriz  $T$  es altamente afectada por el orden de las imágenes que alimentan el nivel más bajo de la Red.

En el presente trabajo manejamos el módulo temporal de una manera diferente poniendo una matriz  $T$  que es una matriz de autocorrelación.

## **Resultados**

En este trabajo se implementó el algoritmo original [George & Jaros] para compararlo contra el optimizado. Las dos redes fueron compiladas con la misma estructura garantizando así la comparabilidad de los resultados experimentales. En ambas versiones se implementó una red de 4 niveles. El nodo ubicado en la cima de la jerarquía es el de clasificación con imágenes de entrada de  $32 \times 32$  píxeles. Así, cada nodo en el nivel más bajo recibe un cuadro de  $4 \times 4$  píxeles que luego se adaptan a los vectores de entrada. Para la validación experimental se construyó un set de datos de pruebas a partir de un modelo esquemático de una fibra en corte longitudinal no medulada de llama.

Para maximizar el rendimiento de una Red HTM la selección de los parámetros para el módulo espacial es de gran importancia. Se debe determinar el umbral  $D$ , como criterio para agregar nuevos centros de cuantización en el módulo espacial y la tasa que finaliza el proceso de aprendizaje. La tasa de aprendizaje se fija en 0.1 y se supone fija para el resto del procedimiento experimental. La distancia  $D$  tanto para el módulo espacial original como el modificado se reguló de manera tal de maximizar la tasa de clasificación. El umbral más apropiado para el método original y el modificado fue de 0.4 y 0.1 respectivamente, alcanzando la tasa de clasificación correspondiente del 74.32% y el 82.33%. Los resultados obtenidos en la comparación de ambas redes han sido resumizados en la Tabla 1.

Tipo Fibra	HTM original [%]	HTM original FGM [%]	HTM original autocorrelac. [%]	Propuesto [%]
No medulada	74.32	74.89	76.34	82.33

Tabla 1. Tasa de clasificación para el HTM original y el propuesto

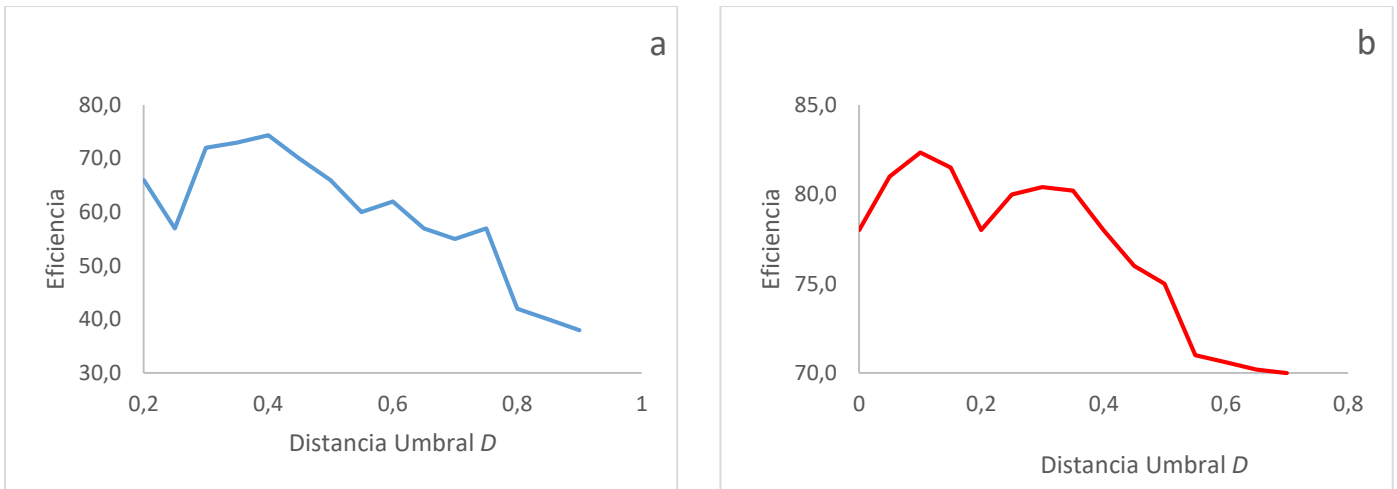


Fig. 1. Eficiencia basada en la distancia umbral  $D$  para el módulo espacial original (a) y el módulo espacial modificado (b).

Se ha presentado un método optimizado para el reconocimiento y clasificación de fibras basado en una Memoria Temporal Jerárquica. El proceso de optimización involucra cambios en los tres diferentes módulos de la Red. Primeramente, las imágenes fueron optimizadas y mapeadas mediante un Filtro de Gabor Modificado de acuerdo a un modelo matemático que permite realzar las características visuales más sobresalientes, aquellas que la vista humana mayormente detecta. A continuación, los centros de cuantización fueron almacenados más uniformemente y, finalmente, el reemplazo del módulo temporal por una matriz de autocorrelación para los centros de cuantización existentes en cada nodo. El procedimiento de aprendizaje se redujo en costo computacional y en consecuencia, decreció el tiempo de entrenamiento. Se logró una eficiencia del 82.33% con la red optimizada contra un 74.32% de la configuración original.

Por otro lado, el presente trabajo intenta contribuir en el proceso de reconocimiento y clasificación de fibra para obtener una medida confiable y soportar indicadores de calidad de la fibra. Además, pretende constituirse en una herramienta útil y accesible que dará respaldo a las investigaciones científicas que se llevan adelante para intervenir en proyectos de desarrollo y promoción humana, y así conducir planes y formular recomendaciones viables para evitar la degradación de los recursos naturales y soslayar problemas de pobreza, marginalidad, emigración y desarraigo entre otros. Finalmente, intenta proporcionar una solución adecuada que permita difundir las cualidades de sustentabilidad para avalar comercialmente los productos textiles de la región.

## Referencias

- Adot, O. (2010) "Introducción a la industrialización de la lana y las fibras especiales", Documento Interno SUPPRAD N°2,
- Bruce, C., Desimone, R. and Gross, C. G. (1981) "Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque", *Journal of Neurophysiology*, 46(2), 369-384.
- Cancio, A., Rebuffi, G., Mueller, J., Duga L. y Rigalt, F. (2016) "Parámetros cualitativos de la producción de fibras de llamas (Lama Gama) machos en la puna argentina", INTA EEA, Bariloche, INTA AER Trancas, INTA EEA Catamarca, Comunicación Técnica, PA 492.
- Ford, J. and Roff, W. (1954) "Identification of textile and related fibers", *Journal of the Textile Institute Transactions*, 45: 580-611.
- Fukushima, k. (1980) "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biological Cybernetics*, 36(4), 193-202.
- George, D. & Jaros, B. (2007) "The htm learning algorithms", Whitepaper, Numenta Inc.

George, D. & Hawkins, J. (2009) "Towards a mathematical theory of cortical microcircuits." PLoS Comput. Biol. 5, e1000532.

George, D. (2008) "How the brain might work: A hierarchical and temporal model for learning and recognition," Ph.D., Department of Electrical Engineering, Stanford University, California, USA

Hawkins, J. and Blakeslee, S. (2004) "On intelligence", 1st ed. New York: Times Books.

Hawkins, J., Ahmad, S. and Dubinsky, D. (2011) "Hierarchical Temporal Memory including HTM Cortical Learning Algorithms", Numenta Inc. White paper.

Huhel, D. H. and Wiesel, T. N. (1962) "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex", Journal de Physiologic, 160 (1), 106-154.

Li, T., Chen, J. (2011) "A fingerprint image enhancement method based on Gabor filter", IEEE Mechanic Automation and Control Engineering (MACE), Second International Conference on 2011, pp 3773 – 3776.

Mittal, A., Pagalthivarthi, K. (2007) "Temporal bayesian network based contextual framework for structured information mining", Pattern Recognition Lett. 28, 1873–1884.

Mueller, J. (2011) "Documento Base de Programa Nacional: Fibras Animales", Instituto Nacional de Tecnología Agropecuaria.

Muller, J. (1993) "Objetivos de mejoramiento genético para rumiantes menores", INTA EEA Bariloche, Comunicación Interna, PA 238.

Yang, J., Liu, L., Jiang, T., Fan, Y. (2002) "A modified Gabor filter design method for fingerprint image enhancement", National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 2002.