

Una nueva Red Neuronal para Clustering y Segmentación basada en el Entorno

Laura Lanzarini* Armando De Giusti†
laural@info.unlp.edu.ar degiusti@info.unlp.edu.ar
Laboratorio de Investigación y Desarrollo en Informática‡
Facultad de Informática. UNLP.

Resumen

Este artículo presenta una nueva red neuronal para clustering y segmentación de los datos o patrones de entrada que, a diferencia de los modelos existentes, no requiere la indicación de parámetros de aceptación dependientes del problema.

En la estrategia propuesta se realiza un preprocesamiento de los patrones a clasificar analizando su relación con sus vecinos más próximos según una medida de similitud. De esta manera, se caracterizan los clusters individualmente, facilitando la aplicación de la red y permitiendo obtener una respuesta específica para cada clase.

Se muestran resultados de su aplicación en problemas cuyas clases poseen diferentes grados de dispersión, así como en imágenes color, correspondientes a muestras histológicas de tejido hepático, sobre las que se realizan tanto tareas de clustering como de segmentación utilizando el mismo método.

La performance de esta nueva propuesta ha sido comparada con una red del tipo winner-take-all (WTA), basada en un único descriptor por clase, ampliamente utilizado en procesos de clustering y con el método CDL que utiliza varios descriptores por clase [2].

Finalmente, se presentan algunas conclusiones así como posibles líneas de trabajo futuras.

1 Introducción

Las técnicas de clustering, como su nombre lo indica, se caracterizan por agrupar los objetos de entrada utilizando alguna medida de similitud. El resultado de este proceso es la formación de *clases o grupos*.

Los elementos a agrupar se encuentran representados a través de sus respectivos vectores de características y se asume que los pertenecientes a una misma clase presentan valores cercanos para una medida de similitud dada.

Esto tiene aplicación en áreas muy diferentes como por ejemplo sonar, radar o diagnóstico médico. Tal amplio rango ha llevado a que la clasificación de patrones no supervisada sea estudiada en detalle [3], [4], [5], [1], [6].

En el área de redes neuronales se han propuesto varias soluciones a este problema.

Hablando específicamente de clustering, las arquitecturas existentes se pueden separar en dos clases: las que utilizan un único representante o descriptor por cada clase, también llamadas winner-take-all y las que utilizan varios descriptores.

Las del primer tipo producen buenos resultados al ser aplicadas a problemas donde las clases presentan muy poca dispersión, ya que con un único representante sólo se puede reconocer lo que pertenece a la hipersfera que lo rodea [9].

Variantes de estos métodos utilizan hipercubos produciendo un efecto similar [8]. En particular los métodos propuestos por Simpson [10], [11] presentan alternativas para clasificar datos en pocas pasadas pero su resultado depende del orden de los patrones de entrada.

*Prof. Adjunto con Dedicación Exclusiva UNLP.

†Prof. Titular Dedicación Exclusiva UNLP.

‡LIDI. Laboratorio de Investigación y Desarrollo en Informática. Facultad de Informática. UNLP. Calle 50 y 115 1^{er} Piso - CP 1900 - La Plata - Bs.As. - Argentina. Tel/FAX:54-221-4227707. E-mail: lidi@info.unlp.edu.ar

Por otro lado, los que utilizan varios representantes, como por ejemplo [2], requieren establecer parámetros iniciales de similitud, dependientes del problema, que permitan relacionar dichos representantes. Los valores de estos parámetros deben ser calculados en base a un subconjunto de representantes de una clase dada y se aplican por igual a todas las clases del problema. Esto impone una fuerte restricción sobre las clases ya que todas deben responder a la misma medida de similitud.

El presente artículo define una nueva red neuronal perteneciente al segundo grupo, que busca mejorar las propuestas anteriores de manera de lograr una clasificación automática que no requiera parámetros iniciales, ni sea dependiente del orden de análisis de los datos y tolere la coexistencia de clases con diferente grado de dispersión.

2 El problema médico

Esta red fue pensada para la clasificación de pixels de imágenes a 256 colores y forma parte de un proceso para el reconocimiento de los elementos de una muestra de tejido hepático actualmente en desarrollo.

Este tipo de muestra está formada en un 60% por hepatocitos cuya identificación y conteo permite detectar diferentes patologías. El 40% restante está formado por vasos sanguíneos, conductillos, tejidos de sostén, y fagocitos especializados.

El objetivo del presente trabajo es lograr aislar los hepatocitos del resto de la muestra para luego separarlos entre sí.

Desde el punto de vista del procesamiento de la imagen se trata de células que no tienen bordes bien definidos con tamaños variables dentro de la misma clase. Este problema, sumado a la diversidad celular presente en una misma muestra, lleva a tener que pensar en la aplicación de métodos no supervisados para ambas etapas.

En esta área, las redes neuronales son una herramienta sumamente útil y en particular, la red CBE aquí propuesta permite solucionar ambos problemas.

Por otro lado, una misma muestra de tejido hepático está formada por aproximadamente 120 imágenes de 640x480 pixels. Tal volumen de datos requiere ser reducido para poder efectuar un procesamiento con un tiempo de respuesta adecuado.

El método aquí planteado se independiza del tamaño de la imagen utilizando el criterio de cardinalidad de cada patrón, es decir, que una imagen de 640x480 pixels puede resumirse en, a lo sumo, 256 patrones; cada uno con su correspondiente cardinalidad (número de repeticiones dentro de la muestra).

Ante este tipo de problema, las soluciones existentes que utilizan un único descriptor por clase sólo permiten obtener hiperesferas o hipercubos alrededor de dicho descriptor. Por otro lado, las que utilizan varios descriptores por clase, asumen que todas las clases tienen la misma medida de dispersión.

La red aquí descrita se basa en el método descripto [7] y en la red neuronal CDL [2], pero mejora la propuesta de esta última incorporando la noción de entorno de cada patrón lo que le permite lograr una respuesta específica para cada clase.

3 Entorno de cada patrón

El criterio utilizado para relacionar los patrones es conexionista. Por este motivo se analiza el entorno de cada uno de ellos a fin de obtener para cada patrón P_i , dos valores: DistMAX y DistMIN.

A continuación se detalla la manera de establecerlos.

Distancia máxima para que dos patrones sean considerados miembros de una misma clase (*DistMAX*)

Sea P_j un patrón ya clasificado como miembro de una clase. P_j , aceptará como *parecidos* a él a aquellos patrones que cumplan

$$dist(P_j, P_i) \leq DistMAX_{P_j} \text{ con } i \neq j.$$

Por lo tanto, todo patrón P_i , con $i \neq j$, que se encuentre a una distancia menor que este valor será considerado similar a P_j y por lo tanto, P_i y P_j tenderán a pertenecer a la misma clase.

Para determinar este valor umbral, se considerarán las tres menores distancias y para cada una de ella se registrará la cantidad de patrones (multiplicada por su cardinalidad). Sea *TotPatrones* la suma de los patrones hallados a estas tres distancias (véase Fig. 3).

$DistMAX_{P_j}$ será la distancia que permita incluir al 50% de *TotPatrones*.

De esta forma, $DistMAX_{P_j}$ será para P_j una medida de proximidad.

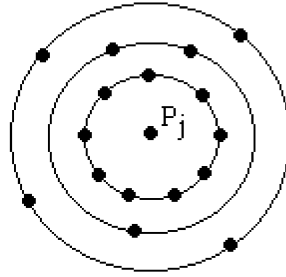


Figura 1: P_j tiene 9 vecinos a una distancia $D1$, 3 a una distancia $D2$ y 4 a una distancia $D3$, con $D1 < D2 < D3$. TotPatrones = 16

En el ejemplo de la Fig. 3, $DistMAX_{P_j} = D1$ y todos aquellos patrones que se encuentren a una distancia menor o igual que $D1$ serán, para P_j , de su misma clase.

El tomar tres distancias mayores que cero forzará al patrón central a tener muchos representantes para poder aislarse, de lo contrario, obtendrá al menos un vecino con quien tenderá a unirse.

Distancia mínima necesaria para ser prototipo ($DistMIN$) Si la distancia entre P_i y P_j es menor que este valor, P_i y P_j serán considerados muy similares, por lo tanto alcanzará con un único descriptor para representarlos a ambos.

Para el patrón P_i , se denotará como $DistMIN_{P_i}$. $DistMIN_{P_j} = \min(dist(P_j, P_i))$ con $j \neq i$

4 Medida de similitud

Para medir la similitud entre un patrón y un prototipo de una clase se utilizará el cuadrado de la distancia euclídea entre ellos y los criterios de aceptación de la clase. Los prototipos son patrones que permiten caracterizar la clase y por lo tanto poseerán los valores $DistMAX$ y $DistMIN$ del punto anterior. La clase obtendrá sus criterios en base a los de cada uno de sus prototipos. Es decir, cada clase poseerá $DistMAX_{Clase j}$ y $DistMIN_{Clase j}$ cuyos valores se obtienen de los prototipos que la forman. Esto se explica con más detalle en el punto 5.

Sean P_i el patrón que se desea identificar y sea T_j el prototipo a analizar representante de la clase S , interesa saber si P_i es similar a T_j . Si esto se verifica, P_i podría pertenecer a la clase S .

Para ser considerados lo suficientemente parecidos como para pertenecer a la misma clase debe verificarse que:

$$\|P_i, T_j\|^2 \leq DistMAX_{Clase S} \quad (1)$$

La ecuación 1 puede escribirse como

$$\|P_i\|^2 + \|T_j\|^2 + 2P_i \cdot T_j \leq DistMAX_{Clase S} \quad (2)$$

$$P_i \cdot T_j + \frac{\|P_i\|^2}{2} \geq \frac{DistMAX_{Clase S} - \|T_j\|^2}{2} \quad (3)$$

Para ser dos prototipos de la misma clase debe cumplirse además que:

$$DistMIN_{Clase S} \leq \|P_i, T_j\|^2 \quad (4)$$

$$\frac{DistMIN_{Clase S} - \|T_j\|^2}{2} \leq P_i \cdot T_j + \frac{\|P_i\|^2}{2} \quad (5)$$

5 Arquitectura de la red

La red neuronal a utilizar consta de tres capas: la capa de entrada, una capa oculta y la capa de salida.

La capa de entrada sólo se encarga de ingresar el patrón.

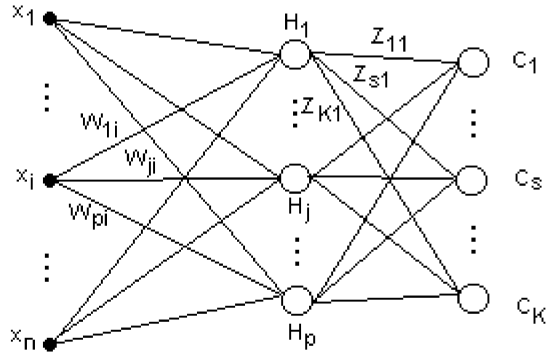


Figura 2: Arquitectura

La capa oculta está formada por perceptrones discretos unipolares. Cada una de estas neuronas contiene la información de un prototipo o descriptor de una clase y el umbral a utilizar está almacenado en la memoria local. El último nivel es equivalente al nivel de Grossberg de la Red de Contrapropagación.

La Fig. 2 muestra la arquitectura de la red que utiliza J prototipos para agrupar los datos de entrada en K clases. La cantidad de prototipos es variable por clase.

6 Aprendizaje

Los patrones de entrada a utilizar son de la forma: $X = (x_1, \dots, x_n)$ formados por los datos originales y agregando información de la norma de la siguiente forma: $X = (x_1, \dots, x_n) = (p_1, \dots, p_m, \|P\|^2/2)$. Es decir que los patrones a utilizar poseen una componente mas que los originales.

El siguiente proceso es aplicado sobre los patrones de entrada y repetido hasta que la definición de las clases se mantenga estable.

Paso 1: Clases iniciales Inicialmente no hay ningún prototipo y por consiguiente ninguna clase asignada.

A partir de este punto, el siguiente proceso iterativo permitirá relacionar los patrones formando las clases correspondientes.

Paso 2 : Tomar el primer patrón de entrenamiento, $X=(x_1, \dots, x_n)$, e ingresarlo a la red.

Paso 3: Medir la similitud con los prototipos existentes:

Calcular la salida de cada uno de los perceptrones de la capa 2 (sólo para aquellos que contengan información de prototipos) de la siguiente forma:

$$S_j = \begin{cases} 1 & \sum_{i=1}^n x_i W_{ji} > U_j \\ 0 & \text{si no} \end{cases} \quad (6)$$

donde

- W_{ji} es el peso del arco que va desde la entrada i hasta el perceptrón j de la capa oculta.
- U_j es el umbral correspondiente al perceptrón j con

$$U_j = \frac{DistMax_{Clase_j} - \|T_j\|^2}{2} \quad (7)$$

Reuniendo la respuesta de todos los perceptrones de la capa intermedia se obtendrá el vector $S = (s_1, \dots, s_j, \dots, s_p)$ donde cada s_j valdrá 1 si X cumple con el criterio de similitud y 0 si no.

Aquí pueden presentarse tres alternativas:

1. $S = (0, \dots, 0)$, es decir, X no cumple el criterio de similitud con ninguno de los prototipos existentes. En este caso, X será considerado el primer prototipo de una clase nueva. Sean P el número de prototipos ingresados hasta el momento y K la cantidad de clases correspondientes. La incorporación

de X permitirá habilitar la neurona (P+1) de la capa intermedia con el siguiente peso para las conexiones que salen de cada una de las neuronas de entrada y llegan a ella:

$$W_{p+1,i} = \begin{cases} x_i & i=1..N \\ 1 & i=N + 1 \end{cases} \quad (8)$$

donde N es la dimensión original del patrón de entrada. Además, para poder calcular posteriormente el valor umbral, se almacenará en la memoria local de H_{P+1} el valor $\|P\|^2$.

En cuanto a la nueva clase, deberá habilitarse la neurona (K+1) de la última capa, C_{K+1} , y asignar los siguientes pesos a cada una de las conexiones que salen de H_{P+1} y llegan a cada una de las neuronas de la capa de salida:

También deberá almacenarse en la memoria local de C_{K+1} , la información del entorno del nuevo prototipo. Esto se reflejará en el valor de $DistMax_{Clase_{K+1}}$ de H_{P+1} .

2. X cumple el criterio de similitud con un único prototipo, es decir que S posee un único elemento igual a 1 y el resto vale 0. Sea J el prototipo que cumple esta condición, es decir $s_J = 1$ y $s_k = 0$ para todo $k \neq J$ y sea S la clase a la que representa. Dado que H_J ya tiene una clase asignada no es necesario habilitar una nueva neurona en la capa de salida.

Si además se cumple que

se habilitará una nueva neurona en la capa oculta asignando los pesos de W como en 5.a) y los de Z de la siguiente forma:

3. Si X cumple el criterio de similitud con varios prototipos puede ocurrir que mas de una neurona de la capa de salida obtenga el valor 1. Esto indica que X pertenece a varias clases. En este caso todas las clases a las que X podría pertenecer se unirán en una sola.

Además, si X cumple la condición (*) con al menos uno de los prototipos existentes, se convertirá en prototipo de la nueva clase.

Paso 4: Borrado de clases pequeñas Se eliminarán todas aquellas clases que contengan menos patrones que el 0.5

Repetir los pasos 3 y 4 hasta que lograr clasificar el 90

7 Aspectos de implementación

- Representación de los patrones de entrada

Es importante tener en cuenta que, según la caracterización que se utilice y el problema de que se trate, los patrones podrían aparecer repetidos, por lo tanto, para cada uno de ellos se tendrá no solo sus características sino su cardinalidad.

- Clases iniciales: Como puede verse, las clases se forman alrededor de los prototipos. Si bien no es necesario contar con clases iniciales, puede acelerarse el proceso de clasificación segmentando el espacio de características en sectores iguales y seleccionando un patrón de cada uno de ellos.

Luego cada patrón será considerado el primer prototipo de una clase nueva.

- Caracterización de clases: Por otro lado, cada prototipo que se agrega a una clase aporta no sólo sus características sino sus valores de similitud DistMAX y DistMIN. La admisión de un patrón como miembro o nuevo prototipo de la clase j estará dada por $DistMAX_{Clase_j}$ y $DistMIN_{Clase_j}$ respectivamente. Cada uno de ellos se obtiene como promedio de los valores de distancia de los prototipos que la forman hasta el momento.

8 Resultados Obtenidos

El método propuesto fue aplicado a dos tipos de problemas diferentes.

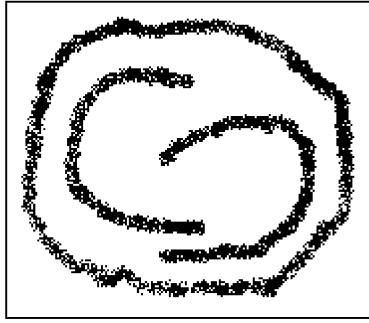


Figura 3: Imagen Original

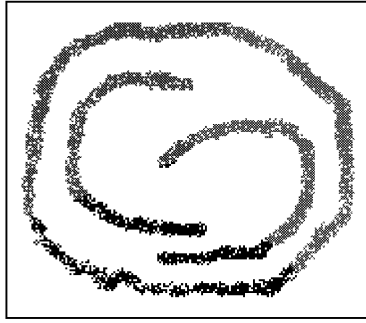


Figura 4: WTA

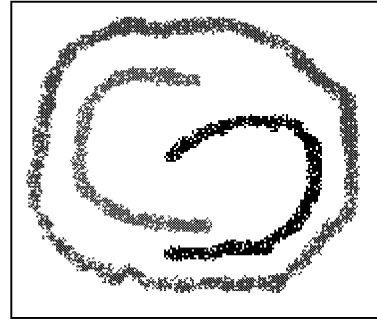


Figura 5: CDL y CBE

8.1 Clases con dispersión alta

Los métodos de clustering convencionales, que utilizan un único descriptor por clase, no son capaces de resolver adecuadamente este problema. Ejemplos de este tipo son el método de las K-Medias (Maravall, 1996) o la redes tipo Kohonen [12], ambos incapaces de realizar una clusterización adecuada.

La posibilidad de disponer de varios descriptores para una misma clase permite generalizar la solución de los métodos antes mencionados. Si bien se aplica la misma medida de similitud con respecto a un descriptor o prototipo determinado, el hecho de disponer de varios de ellos para caracterizar una misma clase permite solucionar el problema de la dispersión. Ejemplos de este tipo son el método CDL [2] y el método CBE aquí descrito.

Como ejemplo de lo antes expuesto se realizará la clasificación de los pixels negros de la imagen binaria de la Fig. 3 caracterizados por su posición. Esto genera un conjunto de patrones sin repetición donde cada clase presenta una dispersión alta.

Como método con un único descriptor, se utilizó un red competitiva. La Fig. 4 demuestra que es incapaz de encontrar la solución correcta.

La Fig. 5 muestra la clasificación realizada mediante los métodos CDL y el propuesto en el presente artículo. Para CDL se asignaron los siguientes parámetros iniciales: $\xi = 0.025$, $\xi_L = 0.035$, $\eta_{min} = 1\%$, $\eta_{max} = 60\%$. La arquitectura aquí propuesta (CBE) logra la clasificación correcta sin ninguna información previa.

8.2 Clases con diferentes grados de dispersión

El punto anterior trata el problema de la dispersión de clases de manera general suponiendo que es una característica de todas ellas. En este punto se trabajará sobre un conjunto de patrones que determina clases con medidas de dispersión muy diferentes entre sí. Esto lleva a obtener tanto clases muy compactas como clases muy dispersas dentro del mismo problema.

El método CBE permite medir la dispersión de cada clase en base a los prototipos incorporados en ella. A medida que se van agregando nuevos descriptores a la clase, su caracterización se va adaptando a ellos. Esto permite obtener, entre clases, diferentes criterios tanto de aceptación y rechazo como de similitud para la definición de nuevos prototipos dentro de la misma.

De esta forma el método CBE permitirá la coexistencia de clases con alta dispersión y clases muy compactas, identificándolas como diferentes.

El método CDL, por su lado, utiliza el mismo criterio de dispersión entre clases. De esta forma, no será factible separar la clases entre sí ya que la medida de similitud es la misma par cada clase.

A modo de ejemplo se realizará nuevamente una clasificación de los puntos negros de la imagen de la Fig. 6 caracterizados por su posición. Puede verse en la figura que hay dos clases muy compactas, las dos clases circulares interiores cuya distancia promedio entre elementos es 1 y otras dos más dispersas; el arco exterior derecho presenta una distancia promedio entre sus elementos de 1.5 y el derecho de 2.3

Para el método CDL no es factible la separación de las 4 clases simultáneamente dado que la medida de similitud es diferente entre una clase y otra. Las figuras 7 y 8 visualizan el resultado de esta arquitectura para diferentes valores de aceptación. Puede verse que para que se consideren los pixels del arco exterior izquierdo como pertenecientes a la misma clase es necesario flexibilizar el criterio, lo que lleva a unirlo con el derecho (Fig. 7). Por otro si se intenta separarlos, el arco izquierdo quedará fraccionado en

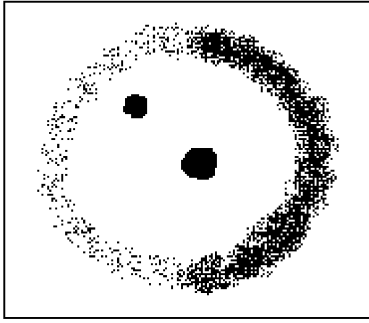


Figura 6: Imagen Original

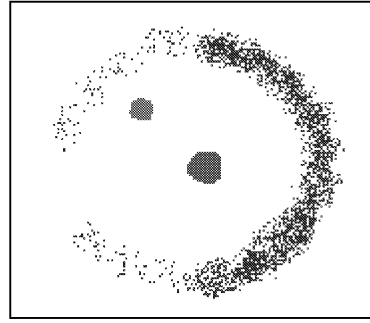


Figura 7: CDL $\xi = 0.01$, $\xi_L = 0.02$, $\eta_{min} = 0.5\%$, $\eta_{max} = 60\%$

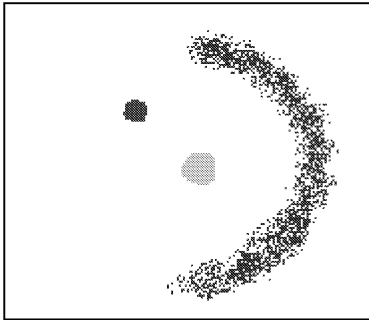


Figura 8: CDL $\xi = 0.01$, $\xi_L = 0.02$, $\eta_{min} = 0.5\%$, $\eta_{max} = 60\%$

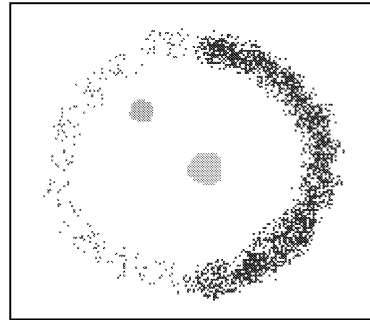


Figura 9: CBE

muchas clases pequeñas y el criterio de eliminación las borrará (Fig. 8). La red CBE no presenta estos inconvenientes ya que tiene la posibilidad de asociar un criterio de aceptación diferente para cada clase (Fig. 9).

9 Ejemplo

Análisis de una muestra histológica (tejido hepático)

A continuación se muestra un ejemplo de clustering y segmentación aplicando la red CBE donde se busca agrupar los pixels de la imagen (Fig. 10) utilizando como característica únicamente su color.

Las Figuras 11, 12 y 13 muestran la aplicación del método de k-medias utilizando dos valores distintos para la cantidad de clases. Dado que la dispersión de las clases a formar es menor que en el ejemplo anterior, el resultado obtenido sólo se ve afectado por el valor de k. Puede verse como la cuarta clase de la aplicación de k-medias con cuatro clases (Fig. 11) se separa en dos (Figs. 12 y 13) al aumentar k mostrando su alta dependencia con el parámetro.

La aplicación del método CDL implica la determinación del valor de los parámetros de entrada. Los valores utilizados en la Fig. 14 fueron obtenidos como resultados intermedios del método CBE. La Fig. 15 es la consecuencia de la aplicación de CDL con parámetros incorrectos.

Finalmente, el método CBE permite arribar la solución de la Fig. 16 sin necesidad de información adicional.

La Fig. 17 muestra la aplicación del método CBE para segmentar los elementos de la Fig. 16. Para obtener este resultado, los pixels han sido caracterizados únicamente por su posición. Si se desea, las clases más pequeñas pueden eliminarse simplemente por su área.

10 Conclusiones

Se ha presentado una nueva red neuronal para clustering y segmentación de los datos o patrones de entrada que, a diferencia de los modelos existentes, no requiere la indicación de parámetros de aceptación

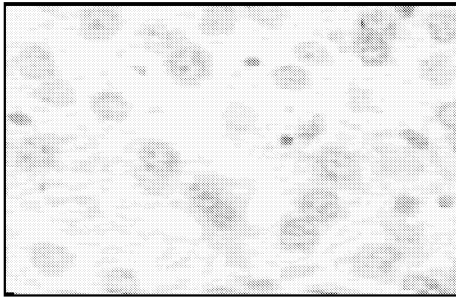


Figura 10: Imagen Original

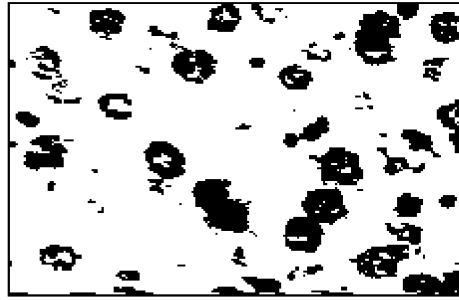


Figura 11: Clase n° 4 de K-Medias con 4 clases ($k=4$)

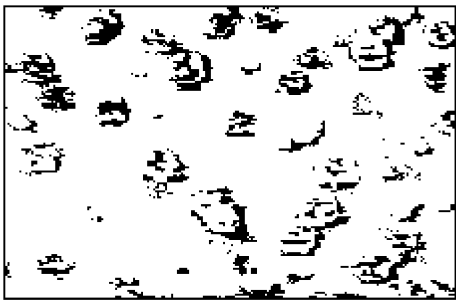


Figura 12: Clase n° 4 de K-Medias con 5 clases ($k=5$)

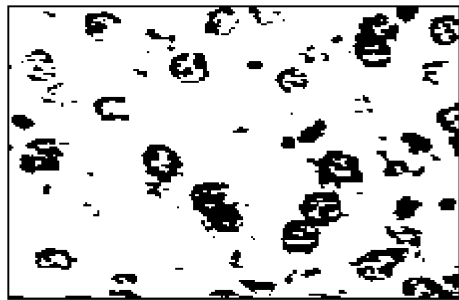


Figura 13: Clase n° 5 de K-Medias con 5 clases ($k=5$)

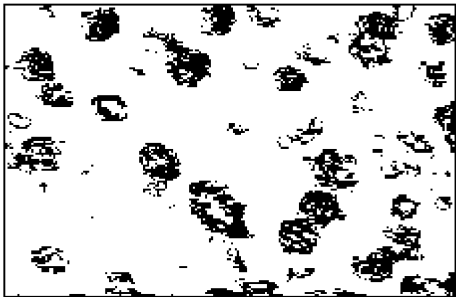


Figura 14: CDL con $\xi = 0.009$, $\xi_L = 0.03$, $\eta_{min} = 0.5\%$, $\eta_{max} = 60\%$

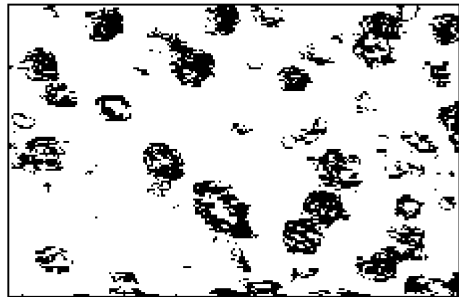


Figura 15: CDL con $\xi = 0.01$, $\xi_L = 0.02$, $\eta_{min} = 0.5\%$, $\eta_{max} = 60\%$

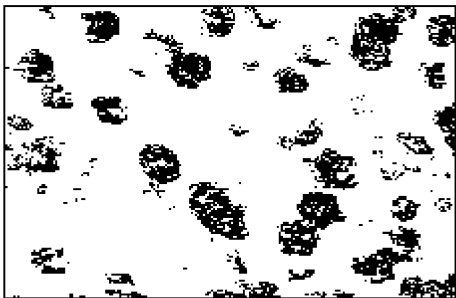


Figura 16: CBE

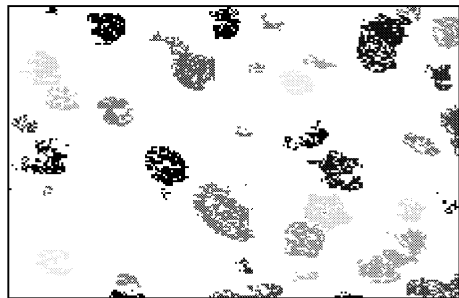


Figura 17: Segmentación

dependientes del problema.

Este método fue pensado para la clasificación de pixels de imágenes a 256 colores y forma parte de un proceso para el reconocimiento de los elementos de una muestra de tejido hepático actualmente en desarrollo. En particular, el método se ha comparado favorablemente con las técnicas clásicas de reconocimiento de patrones.

Si bien los resultados obtenidos son satisfactorios, cuando el tamaño del conjunto de datos a clasificar es grande, la evaluación del entorno de cada patrón puede tener un costo computacional alto. Esto lleva a obtener soluciones dependientes del problema y abre la posibilidad de investigar la paralelización del procesamiento.

En el caso de clustering aplicado a pixels de una imagen caracterizados por su color se ha reducido la cantidad de patrones mediante el concepto de cardinalidad; lo que llevó a un problema de no más 256 patrones. Por otro lado, en el caso de pixels caracterizados por su posición, la cantidad de patrones podrá ser igual al tamaño de la imagen y en este caso deberá buscarse un análisis del entorno que reduzca el tiempo de cálculo.

Pese a lo antes expuesto, la red neuronal aquí propuesta es lo suficientemente general como para aplicarse en ambas situaciones con resultados adecuados.

Toda la documentación se encuentra disponible en el L.I.D.I. (Laboratorio de Investigación y Desarrollo en Informática), 50 y 115 1er. Piso, La Plata.

Referencias

- [1] Jain Anil. *Fundamentals of Digital Image Processing*. Prentice Hall, 1989.
- [2] Torbjorn Eltoft. A new neural network for cluster-detection-and-labeling. *IEEE Transactions on Neural Networks*, 9(5):1021–1035, 1998.
- [3] Maravall Gomez-Allende. *Reconocimiento de Formas y Visión Artificial*. Addison-Wesley Iberoamericana, 1994.
- [4] Gonzalez and Woods. *Digital Image Processing*. Addison-Wesley, 1992.
- [5] Baxes Gregory. *Digital Image Processing*. Wiley, 1994.
- [6] Nicmann H. *Pattern Analysis and Understanding*. Springer Verlag, 1989.
- [7] Lanzarini Laura and De Giusti Armando. Environment based clustering. a new approach. *1st. Int'l Workshop on Image and Signal Processing and Analysis. IEEE CAS and ASSP Croatia*, 2000.
- [8] Massimo Meneganti et al. Fuzzy neural networks for classification and detection of anomalies. *IEEE Transactions on Neural Networks*, 9(5), 1998.
- [9] Newton S et al. Adaptive fuzzy leader clustering of complex data sets in pattern recognition. *IEEE Transactions on Neural Networks*, 3(5):794–800, 1992.
- [10] Patrick Simpson. Fuzzy min-max neural networks-part 1:clustering. *IEEE Transactions on Neural Networks*, 3(5):776–786, 1992.
- [11] Patrick Simpson. Fuzzy min-max neural networks-part 2:clustering. *IEEE Transactions on Fuzzy Systems*, 1(1):32–44, 1993.
- [12] Kohonen Teuvo. *Self-Organizing Maps*. Springer Verlag, 1997.