

Redes neuronales y Series de Fourier aplicados al procesamiento de imágenes en problemas de aprendizaje.

Gabriela Esperón

Universidad de Palermo, Facultad de Ingeniería,
Buenos Aires, Argentina
aigroup@palermo.edu

Alberto E. Laudadio

Universidad de Palermo, Facultad de Ingeniería,
Buenos Aires, Argentina
aigroup@palermo.edu

Alejandro Fiel Martínez

Universidad de Palermo, Facultad de Ingeniería,
Buenos Aires, Argentina
aigroup@palermo.edu

Eduardo Serrano

Universidad de Palermo, Facultad de Ingeniería,
Buenos Aires, Argentina
aigroup@palermo.edu

Daniela López De Luise

**Universidad de Palermo, Facultad de Ingeniería,
Buenos Aires, Argentina
aigroup@palermo.edu**

Abstract

This work is part of the project called PROA which aims to develop an intelligent system for early detection of learning problems for children between 3 and 6 years of age. This document is a description of the results obtained with the implementation of certain Artificial Intelligence strategies to classify and identify plane figures. These figures are obtained from children drawing under test conditions. During the test a child is required to imitate a simple figure, appropriate for its age.

As computing digitalized images (even using just border data) is very expensive, it is interesting to have an approach to reduce the number of original data into a reduced set of parameters.

This is a proposal based on Fourier Fast Transform analysis to reduce data and feed a Neural Network.

Due the regularities of the patterns, the number of reduced coefficients is very small. They could be used as simple classification criteria for such images.

This work focuses in the analysis of some cases and implementations using neural networks called Multilayer Perceptron (MLP) from WEKA framework.

Even thought, graphics evaluation could admit certain subjectivity this new approach could be thought as a new kind of standard criteria.

Keywords: Redes Neuronales. Series de Fourier. Procesamiento de Imágenes

Resumen

El presente trabajo forma parte del proyecto denominado PROA que tiene como objetivo desarrollar un Sistema Inteligente para Detección Precoz de Problemas de Aprendizaje destinado a niños de entre 3 y 6 años de edad.

En este contexto, se presenta una aplicación de técnicas de Inteligencia Artificial a la clasificación e identificación de figuras planas obtenidas a partir de gráficos realizados por infantes en condiciones de test. En el mismo, se requiere al examinado imitar cierta figura sencilla, apta para su edad. En particular, se expone el procesamiento realizado con círculos.

Considerando que el ingreso de las imágenes digitalizadas, o aún de los datos de los bordes, es computacionalmente muy costoso resulta atractivo emplear una estrategia capaz de comprimir los datos originales a una cantidad razonable de parámetros significativos.

Se presenta una alternativa, empleando coeficientes en Series de Fourier para la alimentación y entrenamiento de una red neuronal.

Dado que las imágenes patrón del test son regulares, quedan caracterizadas por un número reducido de coeficientes y pueden clasificarse con criterios sencillos. Por lo tanto, se presenta un escenario favorable para la utilización de Series de Fourier para dicha caracterización.

En este trabajo se analizan casos y diversas aplicaciones empleando una red neuronal de tipo Multilayer Perceptron procesadas con la herramienta WEKA.

Si bien, en la comunidad, puede presentarse cierto grado de subjetividad en la evaluación de los gráficos, se puede presumir con razonable acierto que este tratamiento será la base para la estandarización de ciertos criterios ad hoc.

Palabras claves: Redes Neuronales. Series de Fourier. Procesamiento de Imágenes

INTRODUCCIÓN

El presente trabajo detalla los procedimientos y distintos parámetros considerados en la elaboración y procesamiento de lotes de entrenamiento de una Red Neuronal del tipo Multilayer Perceptron con el fin que aprenda a clasificar si una imagen dibujada por un niño de cierta edad se puede considerar un círculo o no. Esta tarea forma parte del desarrollo de un sistema inteligente para la detección de problemas de aprendizaje **Erreur ! Source du renvoi introuvable.** Para ello se emplearon dos lotes de figuras: el primero se obtuvo de dibujos realizados por niños de entre tres y seis años de edad. El segundo fue generado artificialmente con una función implementada en herramientas informáticas de última generación con el objetivo de poder comparar resultados. La sección I describe las características de los lotes empleados, la elección de materiales y los parámetros utilizados para la digitalización de las imágenes. En la sección II se describen los distintos métodos utilizados para el procesamiento de las imágenes y los resultados obtenidos. En la sección III se describen el entrenamiento de la Red Neuronal realizado con la Herramienta WEKA© y los resultados obtenidos a partir de los distintos lotes. También se expone en este trabajo la comparación de los resultados obtenidos con distintos métodos empleados para la caracterización de imágenes. Finalmente en la Sección IV se formulan las conclusiones y trabajo a futuro.

I. Características de los lotes de imágenes.

Como base de trabajo se utilizaron 93 protocolos de captura sobre la consigna de dibujar un círculo. El lote de imágenes en estudio está compuesto por 43 protocolos correspondientes a niños de 3 años, 31 de 4 años y 19 de 5 años. Según la clasificación de los docentes, los protocolos fueron divididos por edad como se muestra en la **Erreur ! Source du renvoi introuvable.**

Clasificación	Edad		
	3	4	5
NO	23	11	0
SI	20	20	19
Total	43	31	19

Tabla 1 Detalle de los protocolos

Cada protocolo fue clasificado por la correspondiente maestra como “Si” o “No”, según haya cumplido o no la consigna. Paralelamente se codificó una función empleando herramientas de última generación, para generar aleatoriamente círculos irregulares. Como resultado se obtuvieron 5000 círculos que fueron procesados posteriormente de la misma forma que los protocolos de los infantes. Este lote está formado por 2447 gráficos válidos como círculos y 2553 gráficos clasificados como “no círculos” dado que alguno de sus parámetros excedía el margen del 20% de error que se definió en este caso como aceptable. Este lote de círculos generados aleatoriamente, se utilizó como entrada en el entrenamiento de la red, para poder comparar con la muestra obtenida de los niños.

Elección de materiales

El primer problema a resolver fue la elección de los materiales para optimizar el escaneo de los dibujos.

El papel. Para la elección del papel se tuvieron en cuenta distintos factores que inciden en un óptimo escaneo de dibujos.

Luego de varias pruebas de escaneo y dada la naturaleza del trabajo se determinó utilizar papel de 90gr/m². de tipo mate u opaco ya que se comprobó que este tipo es el que permite obtener mejores resultados en la posterior digitalización de las figuras

Los protocolos están realizados en hojas A5 (148mm x 210mm) para limitar la zona de dibujo y facilitar el escaneo de los mismos.

El lápiz. Para poder seleccionar el tipo de lápiz, primero se debe tener claro qué es lo que se dibujará. El factor fundamental a tener en cuenta es la dureza. Los lápices blandos son de mina aceitosa, frágil, oscura y más gruesa. Luego de un análisis previo consistente en la digitalización de dibujos realizados con distintos tipos de lápiz, los de

mina blanda resultaron los más adecuados debido a su gran flexibilidad y expresividad por lo que se eligió emplear una mina tipo 2B.

Digitalización.

Para digitalizar los protocolos se utilizaron scanners comunes, de los que se encuentran en el mercado. Luego de realizar varias pruebas de escaneo, se establecieron los siguientes valores que se muestran en la Tabla 2 como los óptimos respecto de la cantidad de la información y tamaño de archivo de la imagen escaneada.

Parámetros:
Resolución: 50ppp (puntos por pulgada)
Color: 256 bits en escala de grises.
Profundidad: 8 bits
Formato: PNG

Tabla 2. Parámetros de digitalización

Con los parámetros arriba definidos se obtuvieron imágenes de dimensiones: 413x292 pixeles (Ancho x Alto), dando una cantidad de 120596 puntos por imagen.

II. Procesamiento de Imágenes

En esta etapa se buscó procesar las imágenes como datos y prepararlas de forma tal que se pueda alimentar una red neuronal con el objeto de entrenarla y posteriormente utilizar la misma para la clasificación de las figuras.

Para el procesamiento de las imágenes se utilizaron herramientas de última generación que cuentan con un potente motor de procesamiento matemático y capacidad de manipulación de imágenes en formato PNG.

La herramienta empleada permite cargar las imágenes y trabajar con éstas en forma de matriz, facilitándose así la manipulación, reducción y presentación de las mismas.

En una primera etapa se trabajó con dos métodos de compresión de imágenes para estudiar la performance de cada uno y el comportamiento de la Red Neuronal, buscando aquel que resultara más eficiente y eficaz para el entrenamiento de la red y el reconocimiento de figuras.

Inicialmente se trabajó con: Reducción de matriz con filtro no lineal (RFNL) y Reducción de imagen aplicando filtro Hermite. En la **Erreur ! Source du renvoi introuvable.** se presentan los resultados de la red neuronal utilizando RFNL. Con el filtro Hermite, no se consiguió completar ninguna prueba, ya que el tamaño del lote de datos era demasiado grande.

	Cross-validation	Training set
Correctly Classified Instances	69,89%	97,85%
Incorrectly Classified Instances	30,11%	2,15%
Kappa statistic	0,1412	0,9405
Mean absolute error	0,2979	0,0338
Root mean squared error	0,5264	0,1527
Relative absolute error	79,34%	9,02%
Root relative squared error	121,90%	35,39%
Total Number of Instances	93	93

Tabla 3. Resultados MLP con RFNL

Dado los pobres resultados obtenidos, se decidió descartar la compresión de imágenes e intentar con técnicas de caracterización de figuras, empleando Series de Fourier.

Caracterización por Series de Fourier

El primer paso consistió en generar un algoritmo que permitiera calcular el centro y el radio de una figura realizada por los niños, tomando los valores de los ángulos en forma creciente y los radios respectivos a partir de la matriz de unos y ceros correspondiente a cada figura (etiquetada en el algoritmo como *zz*).

Dicho algoritmo se muestra a continuación

```
nz=size (zz) ;
[xi,yi]=find (zz>0) ;
ni=length (xi) ;
mx=mean (xi) ;
my=mean (yi) ;
centro=[mx,my] ;
zc=mx+i*my;
ang=[];
ra=[];
for kk=1:ni,
    zk=(xi (kk)+i*yi (kk)) -zc;
    ang=[ang,angle (zk) ] ;
    ra=[ra,abs (zk) ] ;
end
[angulo,ii]=sort (ang) ;
radio=ra (ii) ;
```

Realizado este procesamiento, se implementó un algoritmo para calcular los coeficientes de Fourier a partir de los radios y ángulos obtenidos con el algoritmo anterior.

Se decidió emplear 32 coeficientes para caracterizar cada figura, ya que este número de elementos de la serie permite conservar las características principales de la figura sin perder información [3][5][4].

Una vez procesadas las imágenes, se obtuvo un archivo de texto por cada imagen, conteniendo estos últimos, en forma de vector binario, la información de las mismas. Éste sería el vector de caracterización de la imagen.

III. Entrenamiento de la Red Neuronal

La Herramienta WEKA © es un conjunto de librerías JAVA para la extracción de conocimiento desde bases de datos. Al estar programado en Java, es independiente de la arquitectura, y funciona en cualquier plataforma sobre la que haya una máquina virtual Java disponible.

Con este software se realizó el entrenamiento y prueba de la red neuronal[6].

Formateado de vectores.

Antes de alimentar la red, se deben procesar los archivos que contienen los vectores de características de cada imagen, presentándolos en un formato de datos apto para la herramienta empleada. Para ello se utilizó un script desarrollado en Python que prepara el encabezado requerido junto con los vectores en un mismo archivo para que pueda ser ingresado a la red.

Una vez obtenido el archivo en formato adecuado para WEKA©, se procedió a la carga de datos y correspondiente entrenamiento de la Red Neuronal.

Las pruebas se realizaron con dos métodos de la función clasificadora Multilayer Perceptron:

Use training set: Con esta opción Weka entrenará el modelo con todos los datos disponibles y a posteriori realizará la evaluación sobre los mismos datos.

Cross-validation: Se realiza la evaluación mediante la técnica de validación cruzada. En este caso se puede establecer el número de particiones a utilizar. [6].

Las tablas 4 y 5 muestran los resultados del entrenamiento de la red neuronal realizado a partir de los protocolos obtenidos de los infantes.

Use training set (muy optimista)

Correctly Classified Instances	100,00%
Incorrectly Classified Instances	0,00%
Kappa statistic	1,0000
Mean absolute error	0,0203
Root mean squared error	0,0298
Relative absolute error	4,37%
Root relative squared error	6,19%
Total Number of Instances	93

Tabla 4

10-fold cross-validation (muy pesimista)

Correctly Classified Instances	68,82%
Incorrectly Classified Instances	31,18%
Kappa statistic	0,3236
Mean absolute error	0,3327
Root mean squared error	0,5146
Relative absolute error	71,54%
Root relative squared error	106,77%
Total Number of Instances	93

Tabla 5

Paralelamente, se realizó el entrenamiento de la red con el lote de 5000 círculos generados aleatoriamente, obteniéndose en este caso resultados significativamente mejores como puede observarse en la tablas 6 y 7.

Use training set (muy optimista)

Correctly Classified Instances	98,90%
Incorrectly Classified Instances	1,10%
Kappa statistic	0,9780
Mean absolute error	0,0148
Root mean squared error	0,1055
Relative absolute error	2,95%
Root relative squared error	21,10%
Total Number of Instances	5000

Tabla 6

10-fold cross-validation (muy pesimista)

Correctly Classified Instances	85,90%
Incorrectly Classified Instances	14,10%
Kappa statistic	0,7183
Mean absolute error	0,1424
Root mean squared error	0,3544
Relative absolute error	28,51%
Root relative squared error	70,90%
Total Number of Instances	5000

Tabla 7

El modelo obtenido a partir de este último entrenamiento con el lote de círculos generados aleatoriamente se utilizó para clasificar el lote de imágenes realizadas por los niños.

Los resultados se muestran en la tabla **Erreur ! Source du renvoi introuvable.**

Supplied test set

Correctly Classified Instances	100,00%
Incorrectly Classified Instances	0,00%
Kappa statistic	1,0000
Mean absolute error	0,0203
Root mean squared error	0,0298
Relative absolute error	4,37%
Root relative squared error	6,19%
Total Number of Instances	93

Tabla 8. Resultados de la clasificación de las imágenes de los infantes.

Como puede observarse se obtuvo un valor del estadístico Kappa igual a 1. Esto indicaría la existencia de problemas en el entrenamiento de la red.

Las causas de estos problemas podrían ser:

A) Mecanismo de aprendizaje.

La red fue entrenada para identificar distintas distorsiones sin tomar en cuenta la edad. Cuando se le presenta una muestra con diferentes edades, "acomoda" lo que detecta al aprendizaje que realizó: Cualquier imagen dentro de un rango que sea compatible con la deformación típica de cualquier edad se clasifica como un círculo.

Por lo tanto habría que considerar realizar el entrenamiento de la red tomando en cuenta la edad lo cual podría reducir la tolerancia de las deformaciones en ciertos casos.

B) Tamaño de la población.

El número de instancias suministradas a la red no resulta suficiente para que esta discrimine grandes deformaciones. A consecuencia de esto nunca llega a determinar un caso de "no círculo". Se puede inferir la presencia de un sobre ajuste del aprendizaje sobre la muestra.

Este problema podría resolverse entrenando la red con más instancias e ingresando deformaciones superiores al 20% para que la red las discrimine.

Con el propósito de generar una nueva muestra se realizó un análisis de la población empleando dos algoritmos de agrupamiento de datos.

B.1) Simple K means: se empleó para detectar la naturaleza de los conjuntos de parámetros de cada instancia, considerando siempre la clasificación original.

Servirá para validar si la muestra se puede clasificar naturalmente en los $k=3$ clusters de edad usando sólo los valores de los parámetros.

Al realizarlo se conforman 3 clusters (Tabla 9) mucho más homogéneos en comparación con la muestra real como se observa en la **Erreur ! Source du renvoi introuvable.1**.

0	37 (40%)
1	30 (32%)
2	26 (28%)

Tabla 9. Resultados utilizando K Means.

Analizando las diferencias poblacionales se podría inferir que se trata de una población muy diversa que se comporta de manera muy homogénea, no siendo naturalmente divisible por edades. Esto indica que no existiría una diferencia significativa por edad en la producción de círculos.

También podría atribuirse al bajo grado de dificultad diferenciadora de la figura a reproducir. Resultaría interesante estudiar a futuro si en ciertas patologías severas estas características se presentan como outliers (valores atípicos).

B.2) Expectation Maximization: Desde el punto de vista de las similitudes poblacionales se utilizó el algoritmo de agrupamiento denominado Expectation Maximization (EM) para averiguar cuáles son las características de la muestra. Los resultados se exponen en la **Erreur ! Source du renvoi introuvable.**

0	3 (3%)
1	10 (11%)
2	11 (12%)
3	15 (16%)
4	18 (19%)
5	10 (11%)
6	13 (14%)
7	8 (9%)
8	5 (5%)

Tabla 6. Resultados utilizando EM.

Se puede observar que la muestra se comporta como una población normal no sesgada. Quedaría pendiente realizar una prueba de normalidad para confirmarla.

En este punto resultaría conveniente utilizar muestras discriminadas por edad para entrenar la red neuronal.

IV. Conclusiones

Si bien los resultados obtenidos a partir de los protocolos de los niños no satisfacen las expectativas, en principio señalan la necesidad de separar los mismos por edades.

Debido a la edad de la población objetivo a la que está dirigido el proyecto, se presentan numerosas dificultades para la obtención de datos experimentales. Ello obligó a trabajar con una muestra relativamente pequeña, en relación a las diferencias que se presentan en los dibujos de los niños de 3 a 6 años. La performance de la red neuronal se vio alterada por esta dificultad.

Para mejorar los resultados, sería necesario tomar cada rango de edad por separado, buscar la unificación de criterios con los docentes y ampliar el tamaño de la muestra, para contar con un lote lo suficientemente representativo que permita respaldarlos sólidamente .

En el caso del entrenamiento realizado con el lote de círculos generados aleatoriamente, los resultados estarían indicando un sobre aprendizaje, lo que provocaría que la red no esté logrando distinguir entre un círculo y un no círculo. Se advierte, en principio, que se debería mejorar el entrenamiento mediante la segmentación del mismo por edades.

Del mismo modo, aún resta por analizar los protocolos por edades y verificar si los procedimientos empleados son transferibles a la clasificación de otras figuras.

Sólo en este caso la herramienta que se propone, podría resultar un aporte significativo a la comunidad educativa como uno de los módulos componentes del sistema de detección temprana PROA.

Referencias

- [1] Fiel Martínez, A., Esperón G., López De Luise, D., Laudadio, E., Burgos, M. Sistema para la detección precoz de problemas en el aprendizaje. En Anales del WICC 2008 (Mayo 2008), pp. 10-14.
- [2] Hertz, J., Krogh, A. y Palmer, R: Introduction to the theory of neural computation. Decimo segunda edición, Redwood City: Addison-Wesley. 1996
- [3] Duoandikoetxea, J. Lecciones sobre las series y transformadas de Fourier. UNAN-Managua, 2003
- [4] Vélez Serrano et al., Visión por computador. 2003.
- [5] Series de Fourier Transformadas de Fourier y Aplicaciones. En Divulgaciones Matemáticas v. 5, No. 1/2 (1997), pp. 43-60
- [6] Witten, I. y Frank, E. Data Mining. Practical Machine Learning Tools and Techniques. San Francisco: Morgan Kaufmann Publishers. 2005