

Compresión de Imágenes para Diagnóstico Médico utilizando Redes Neuronales

Lic. Lanzarini Laura¹

A.C. María Teresa Vargas Camacho²

Dr. Amado Badrán³

Ing. De Giusti Armando.⁴

Laboratorio de Investigación y Desarrollo en Informática.⁵
Departamento de Informática - Facultad de Ciencias Exactas.
Universidad Nacional de La Plata

Resumen

La compresión de imágenes es un tema ampliamente estudiado. Las soluciones convencionales ofrecen radios de compresión variables en función de la imagen de la cual se trate y en general, no ofrecen buenos resultados para imágenes con muchas variaciones en sus tonalidades

El presente trabajo es una aplicación de compresión de imágenes de tomografías computadas de pacientes utilizando redes neuronales, que permite realizar tanto la compresión como la descompresión de las imágenes con un radio fijo de 8:1 y con una pérdida del 2%.

Palabras Claves: Inteligencia Artificial - Redes Neuronales - Procesamiento de Imágenes - Diagnóstico médico.

Introducción.

La mayoría de los fenómenos que se estudian en medicina, debido a procesos tanto normales como patológicos, tienen fundamentos morfológicos que se expresan por imágenes capaces de definir el grado o intensidad del fenómeno, así como sus variaciones temporales.

El avance tecnológico ha facilitado el desarrollo de nuevos campos de la medicina como es el del diagnóstico por imágenes, que posibilitan un estudio más completo del paciente por métodos no invasivos.

¹ Prof. Adjunto dedicación Exclusiva del Dpto. de Informática, Fac. Cs. Exactas, UNLP.

² Auxiliar Docente dedicación Simple del Dpto. de Informática, Fac. Cs. Exactas, UNLP.

³ Prof. Titular dedicación Exclusiva del Dpto de Ciencias Médicas, UNLP.

⁴ Director del LIDI. Investigador Principal CONICET. Prof. Titular dedicación Exclusiva del Dpto. de Informática, Fac Cs. Exactas, UNLP.

⁵ LIDI. Laboratorio de Investigación y Desarrollo en Informática, Dpto. de Informática, Fac. de Cs. Exactas, UNLP. Calle 50 y 115 - 1^{er} piso - 1900 La Plata - Buenos Aires - Argentina.
Tel / Fax: 54 - 21 - 22 7707
E-Mail: lidi @ ada.info.unlp.edu.ar

Estos métodos incluyen desde una radiografía (RX), hasta técnicas más complejas como la ecografía, tomografía axial computada (TAC), resonancia magnética (RM), etc. todas ellas en forma directa o utilizando sustancias de contraste.

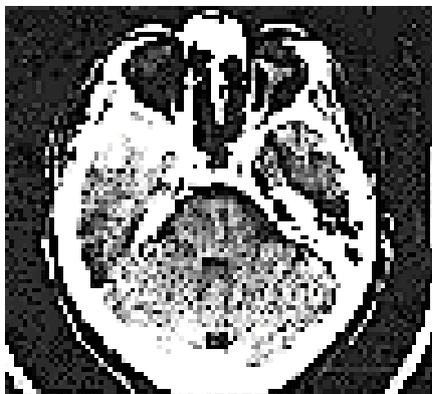
El estudio de una única imagen, permite obtener datos valiosos tanto para el diagnóstico como para algunos tratamientos: una RX dental permite diagnosticar lesiones y adoptar un tratamiento específico (extracción de una pieza dentaria) y hasta la identificación de ciertos individuos. Una TAC de la extremidad cefálica, permite el diagnóstico de una lesión o de un patología craneana o cerebral y establecer un tratamiento adecuado.

En los casos citados, los diagnósticos se realizan por variaciones cualitativas o cuantitativas en los valores de ciertos parámetros que incluyen una gran variedad de mediciones: distancia entre elementos anatómicos, formas, tamaños y amplitud de algunos órganos o lesiones y sus relaciones con estructuras vecinas. En base al análisis de las imágenes el médico establece un probable diagnóstico que le permite inferir un pronóstico y establecer un tratamiento.

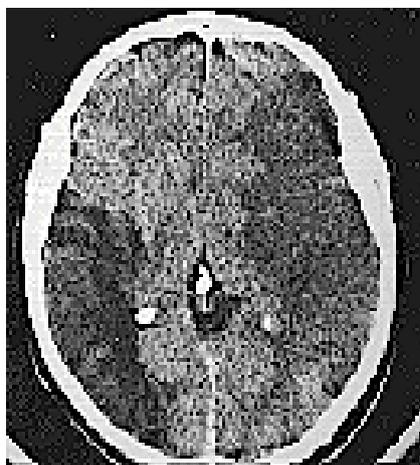
Estos importantes beneficios que brinda el diagnóstico por imágenes, se complementa con la enorme importancia que tiene el método para el estudio evolutivo de las lesiones o patologías diagnosticadas.

De esta manera es posible cuantificar la magnitud de los cambios así como la velocidad con que se producen, permitiendo establecer tanto el grado de malignidad como la gravedad del proceso o la efectividad de un tratamiento.

Este estudio evolutivo sólo es posible comparando imágenes tomadas en distintos tiempos, para lo cual es necesario contar con un sistema de archivo de imágenes adecuado.



En el ejemplo que se adjunta al presente trabajo se puede definir la localización de los planos de corte en que se realizan mediciones y obtener beneficios de tales mediciones como son: determinar el grosor del hueso de la calota craneana o de cualquier otro elemento óseo identificable, la densidad ósea, la forma y tamaño de los ventrículos cerebrales laterales, la densidad de la masa cefálica, la forma disposición y dirección del nervio óptico, etc.



La obtención de imágenes similares del mismo individuo en distintas edades permitiría la comparación de dichos parámetros (desarrollo ontogenético) o el análisis de una lesión patológica evolutiva; la disponibilidad de imágenes de otros individuos, permitirían estudios filogenéticos o de poblaciones definidas.

En cualquiera de estos casos es imprescindible poder recuperar las imágenes archivadas con el máximo de exactitud y reproducibilidad, ya que la comparación de los valores de los parámetros utilizados, permitirán establecer correlaciones.

Objetivo:

Se plantea la utilización del modelo de red neuronal backpropagation aplicado a compresión de imágenes de tomografías computadas, formado por tres capas. Dicho modelo utiliza entrenamiento supervisado basado en aprendizaje por corrección del error entre la salida esperada y la obtenida.

Dado que se busca obtener un algoritmo que permita comprimir y descomprimir una imagen, se coloca en los extremos de la red el patrón de entrada y se utilizan las neuronas de la capa oculta para representar la compresión.

Esto implica que una vez superada la etapa de entrenamiento, el uso de la red consiste en duplicar la capa oculta y utilizar

- la capa de entrada y la oculta para realizar el proceso de compresión y
- la capa oculta y la de salida para realizar la descompresión.

Análisis de la Red Neuronal utilizada:

Red Backpropagation:

En este trabajo se utilizó una red de propagación hacia atrás (Backpropagation Network) la cual está diseñada para funcionar como red multicapa, empleando un método de aprendizaje supervisado.

Características de la Red Backpropagation utilizada:

Se consideraron tres capas, la de entrada, la de salida y una capa oculta. Cada neurona está conectada a todas las neuronas de la capa siguiente.

Capa de Salida

...

...

...

Capa de entrada

La red backpropagation es una red de correspondencia, es decir, permite calcular una relación funcional entre la entrada y la salida. A cada patrón de entrada se le asocia su patrón de salida correspondiente.

Entrenamiento:

La red aprende un conjunto predefinido de pares de entradas y salidas empleando un ciclo de propagación-adaptación de dos fases:

La trama de entrada se aplica a la primera capa y se propaga a través de las capas superiores hasta llegar a la capa de salida. La señal de salida se compara con la salida deseada y se calcula una señal de error para cada unidad de la capa de salida.

Las señales de salida se transmiten hacia atrás, en dirección a los nodos de la capa intermedia que contribuyen directamente sobre la capa de salida.

Los nodos de la capa intermedia reciben una fracción del error en forma proporcional a su contribución en el valor de salida.

Basándose en la señal de error percibida, se actualizan los pesos de conexión de cada unidad para hacer que la red converja hacia un estado que permita codificar todas las tramas de entrenamiento.

Las ecuaciones utilizadas se resumen a continuación:

Se aplica el patrón de entrada $X_p = (x_{p1}, x_{p2}, \dots, x_{pN})$ a las unidades de entrada.

Se calculan las entradas netas para la capa oculta

$$\text{neta}^h_{pj} = \sum_{i=1}^N w^h_{ji} X_{pi} + \theta^h_j$$

Se calculan las salidas para la capa oculta $i_{pj} = f^h_j(\text{neta}^h_{pj})$

Se calculan los valores netos de las neuronas de la capa de salida

$$\text{neta}^o_{pk} = \sum_{j=1}^L w^o_{kj} i_{pj} + \theta^o_k$$

Se calculan las salidas $o_{pk} = f^o_k(\text{neta}^o_{pk})$

Se calculan los términos de error para las unidades de salida:

$$\delta^o_{pk} = (y_{pk} - o_{pk}) f'^o_k(\text{neta}^o_{pk})$$

Se calculan los términos de error para las unidades ocultas:

$$\delta^h_{pj} = f'^h_j(\text{neta}^h_{pj}) \sum_k \delta^o_{pk} w^o_{kj}$$

Se actualizan los pesos de la capa de salida y los de la capa oculta

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \eta \delta_{pk}^o i_{pj}$$

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \eta \delta_{pj}^h x_i$$

Aspectos de Implementación:

Para trabajar se utilizaron imágenes en tonos de grises correspondientes a tomografías computadas de diferentes pacientes. Las imágenes son captadas en tonos de grises.

Si bien este trabajo no utiliza imágenes en colores, las ideas aquí utilizadas pueden aplicarse de la misma forma.

Sin embargo, es importante tener en cuenta que cada neurona de la capa de entrada recibe un valor proporcional al de un pixel de la imagen. Es decir que si el pixel tiene color, puede llevar asociado $256 * 3$ valores distintos. Esto implica tener que disponer de un conjunto de patrones de entrenamiento mucho mayor para imágenes en color que para las que son en tonos de gris, lo que representa un mayor tiempo de entrenamiento ya que existen muchas más combinaciones posibles.

En este caso, se utilizó una red con 64 neuronas en la capa de entrada, 64 en la de salida y 8 en la capa oculta.

La función de activación utilizada fue $f(x) = 2 / (1 + \exp(-x)) - 1$ que es una función derivable y permite manejar valores en $(-1..1)$

El color de cada pixel fue llevado de $[0..255]$ a $[-1..1]$, mediante una conversión lineal, a fin de ser utilizado como valor para la capa de entrada. Esto favorece el tiempo de convergencia, ya que de esta forma, se aprovechan los cambios de la función sigmoide cerca del origen.

Los pesos de las conexiones entre las neuronas fueron inicializados aleatoriamente con valores random entre $[-1..1]$.

Dado que el aprendizaje de la red es supervisado, se utilizaron como salida deseadas los mismos valores de la capa de entrada. Es decir, que la compresión de la imagen se ve reflejada por la capa intermedia.

La imagen fue ingresada a la red por filas, tomando 64 pixels por cada vez. Esto permite obtener varios patrones de entrenamiento con una única imagen. Cabe considerar que, no sería conveniente aumentar demasiado el tamaño de la capa de entrada ya que si bien esto puede reducir el tiempo de procesamiento de la imagen, aumenta la cantidad de combinaciones posibles para formar dicho patrón, incrementando el tiempo de convergencia.

El objetivo del entrenamiento de la red es llevar los pesos de las conexiones entre neuronas hacia valores que minimicen el error que representa la diferencia entre el patrón de entrada y el de salida.

Es decir que se minimiza $\sqrt{\Sigma}$ (valor deseado - valor obtenido)² hasta alcanzar la tolerancia deseada. Dicho valor está en relación directa con la calidad de la imagen a obtener, ya que aumentando el valor de la tolerancia del error se obtendrá un menor tiempo de convergencia así como una mayor pérdida en la imagen.

Asimismo, es importante destacar que el radio de compresión del método propuesto en este trabajo es independiente de la imagen de la cual se trate ya que queda perfectamente determinado en base al tamaño de la capa oculta.

Resultados obtenidos:

Inicialmente se consideró una red con 16 neuronas en su capa intermedia, 64 en la de entrada y 64 en la de salida, es decir con una compresión 4:1. Posteriormente se utilizó la técnica de pruning hasta lograr reducir la capa intermedia a 8 neuronas, es decir, logrando un radio de compresión 8:1 sin afectar la tolerancia buscada, ni incrementar demasiado el tiempo de convergencia.

Las pruebas realizadas permitieron obtener un error de 0.01 luego de entrenar a la red con imágenes en tonos de grises, de 220 x 240 pixels realizando 50 iteraciones sobre cada imagen.

El uso de imágenes representativas como forma de obtener los patrones de entrenamiento permite reducir el tiempo de convergencia. Sin embargo, este proceso implica que si se cambia o aumenta el tipo de las imágenes a comprimir, será necesario reentrenar.

El valor de η utilizado fue de 0.01, ya que las pruebas realizadas detectaron que un valor mayor podría producir oscilaciones en la convergencia de la red, incrementando el tiempo de entrenamiento.

En este caso el parámetro de tendencia no fue utilizado, ya que las pruebas realizadas demostraron que no era necesario.

Conclusiones:

Se ha presentado una solución a la compresión de imágenes de tomografías computadas de pacientes, utilizando redes neuronales.

Si bien la red Backpropagation está pensada para establecer una correspondencia entre la entrada y la salida, vemos en este trabajo una aplicación clara del uso de la capa oculta.

A diferencia de los métodos de compresión convencionales, este algoritmo no utiliza directamente los pixels redundantes como un punto clave de la compresión sino que, los resultados a obtener dependen en gran medida de los patrones utilizados en el entrenamiento.

Por la naturaleza de las redes neuronales, este algoritmo es directamente paralelizable ya que las neuronas de una misma capa trabajan de manera totalmente independiente, con lo cual puede adaptarse fácilmente a una arquitectura paralelo.

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.

Bibliografía:

- [BAX] "Digital Image Processing". Gregory A. Baxes. De. Wiley.1994
- [GON] "Digital Image Processing". Rafael González. Addison Wesley.1992
- [JAI] "Fundamentals of Digital Image Processing". Anil Jain. Prentice Hall. 1989
- [KAY] "Graphics File Formats". Kay & Levine.1992
- [KOS] "Neural Networks and Fuzzy Systems". Bart Kosko. Prentice Hall. 1992
- [MAS] "Signal and Image Processing with Neural Networks". Timothy Masters.Wiley & Sons.
1994
- [PAO] "Adaptative Pattern Recognition and Neural Networks" Yoh-Han Pao.Addisson
Wesley. 1989
- [RAO] "Neural Networks and Fuzzy Logic". Rao y Rao. MIS Press. 1995
- [ROS] Ecuaciones Diferenciales. Shepley Ross. Editorial Reverte. 1989