

Red neuronal artificial para el reconocimiento de emociones humanas bajo un algoritmo de retropropagación

Alexander Caicedo

<https://orcid.org/0000-0001-8729-5793>
alex.caicedoalcivar@gmail.com
Banco General Rumiñahui
Quito, Ecuador.

Anthony Caicedo

<https://orcid.org/0000-0001-7550-4754>
anthony.caicedo@pucese.edu.ec
Public Promueve
Quito, Ecuador

Recibido (10/07/21), Aceptado (15/08/21)

Resumen: La era de la revolución tecnológica impulsa cada vez más el desarrollo de tecnologías que facilitan de una u otra manera actividades de la cotidianidad de las personas, generando así un gran avance en el procesamiento de información. Este trabajo tiene como finalidad implementar una red neuronal que permita clasificar los estados emocionales de una persona partiendo de las distintas gestualidades humanas. Se utiliza una base de datos con información de estudiantes de la escuela de Ingeniería en Sistemas y Computación de la PUCE-E. Dicha información son imágenes que expresan las gestualidades de los alumnos y con las cuales se lleva a cabo el análisis comparativo con los datos de entrada. El entorno en el cual converge este trabajo propone que la implementación de este proyecto se realice bajo la programación de una red neuronal de tipo multicapa. Las redes neuronales de alimentación de múltiples capas poseen una serie de propiedades que las hacen particularmente adecuadas para problemas complejos de clasificación de patrones. Back-Propagation, que es un algoritmo de retro propagación empleado en la red neuronal Feedforward fue tomado en consideración para resolver la clasificación de las emociones.

Palabras Clave: Procesamiento de imágenes, redes neuronales, gestualidad, back-propagation, feedforward, clasificación, emociones.

Artificial neural network for the recognition of human emotions under a backpropagation algorithm

Abstract: The era of the technological revolution increasingly encourages the development of technologies that facilitate in one way or another people's daily activities, thus generating a great advance in information processing. The purpose of this work is to implement a neural network that allows classifying the emotional states of a person based on the different human gestures. A database is used with information on students from the PUCE-E School of Computer Science and Engineering. Said information are images that express the gestures of the students and with which the comparative analysis with the input data is carried out. The environment in which this work converges proposes that the implementation of this project be carried out under the programming of a multilayer neural network. Multilayer feeding neural networks possess a number of properties that make them particularly suitable for complex pattern classification problems [8]. Back-Propagation [4], which is a backpropagation algorithm used in the Feedforward neural network, was taken into consideration to solve the classification of emotions.

Keywords: Image processing, neural networks, gestures, back-propagation, feedforward, classification, emotions.



I. INTRODUCCIÓN

La evolución en el desarrollo de software junto con la electrónica ha permitido construir nuevas tecnologías, aplicaciones y dispositivos que son manejados a través de la gestualidad humana, dando respuestas a procesos concretos.

En este trabajo se desarrolló un sistema inteligente que caracteriza los rasgos faciales para conocer las emociones humanas. Para el desarrollo de este sistema se utilizaron 68 imágenes de diferentes características emotivas que diferencian las emociones.

Se empleó el software de programación de Matlab 2015 para el procesamiento y desarrollo de la red neuronal, ya que permite el manejo óptimo de los recursos y el procesamiento de los datos de forma agilizada.

Una red neuronal artificial es una técnica de inteligencia artificial que tiene la capacidad de aprender de las experiencias, mejorando su rendimiento al adaptarse a los cambios en el entorno [11].

Las redes neuronales profundas (DNN) son modelos potentes que han logrado un rendimiento excelente en tareas de aprendizaje difíciles. Aunque los DNN funcionan bien cuando hay disponibles conjuntos de entrenamiento etiquetados grandes, no se pueden usar para asignar secuencias a las secuencias [9]. En los últimos años, las redes neuronales artificiales profundas (incluidas las recurrentes) han ganado numerosos concursos de reconocimiento de patrones y aprendizaje automático [10].

Existe un aspecto fundamental que los diseñadores de redes neuronales enfrentan actualmente y no es más que es elegir un tamaño de red apropiado para una aplicación determinada [6]. El tamaño de la red implica en el caso de las arquitecturas de redes neuronales en capas, el número de capas en una red, el número de nodos por capa y el número de conexiones.

Está claro que la velocidad de aprendizaje de las redes neuronales feedforward es en general mucho más lenta de lo requerido y ha sido un cuello de botella importante en sus aplicaciones durante las últimas décadas [7]. Dos razones principales detrás pueden ser: 1. Los algoritmos de aprendizaje basados en gradiente lentos son ampliamente utilizados para entrenar las redes neuronales, y 2. Todos los parámetros de las redes están sintonizados de forma iterativa mediante el uso de este tipo de algoritmos de aprendizaje

II. DESARROLLO

En este trabajo se empleó el procesamiento de imágenes utilizando redes neuronales artificiales con el uso del algoritmo Back-Propagation [5]. Esta red neuronal se encuentra dentro del grupo de red de aprendizaje supervisado el cual trabajan bajo un gradiente descendente, donde el patrón de entrada se propaga por toda la red hasta la salida, así mismo las salidas de error se propagan hacia la capa anterior de neuronas [2].

Para esta investigación se trabajó con 68 imágenes diferenciadas según la emoción que presenten cada uno de los 17 individuos que conformaron este estudio. Se diferenciaron cuatro gestualidades relacionadas con las siguientes emociones: ira, alegría, tristeza y susto; es decir, por cada individuo del estudio se capturaron cuatro imágenes. Se empleó el software de Matlab 2015 con el toolbox de redes neuronales, ya que ofrece una amplia variedad de comandos y versatilidad para la ejecución del sistema y la obtención de resultados óptimos.

III. METODOLOGÍA

La aplicación desarrollada consiste en analizar los rostros de los estudiantes del 8vo, 9no semestres de la Escuela de Sistemas y Computación de la universidad Pontificia Universidad Católica Sede Esmeraldas (PUCE-SE). Esta permite ingresar el rostro de los estudiantes a través de una imagen para ser analizados con la base de datos y así determinar cuál es el estado emotivo del estudiante; el procesamiento de la imagen se realizó primero redimensionando la imagen y convirtiendo la misma a escala de grises. En las imágenes en escala de gris, los componentes R, G y B tienen la misma intensidad en el espacio RGB [3]. Pero, son más fáciles de procesar que las imágenes a color.

A partir de las imágenes en escala de grises se extrajeron los ojos y las porciones de la boca ya que estas áreas contienen la información más esencial de la emoción. Estas áreas extraídas se contabilizaron en una matriz de 64X64 cada una [1]. Estas matrices contienen valores numéricos que van de 0 a 255, las mismas que se manipularon para la extracción de características.

En la figura 1 se presenta el esquema del sistema desarrollado donde se observan cinco etapas secuenciales que caracterizan el proceso de clasificación.

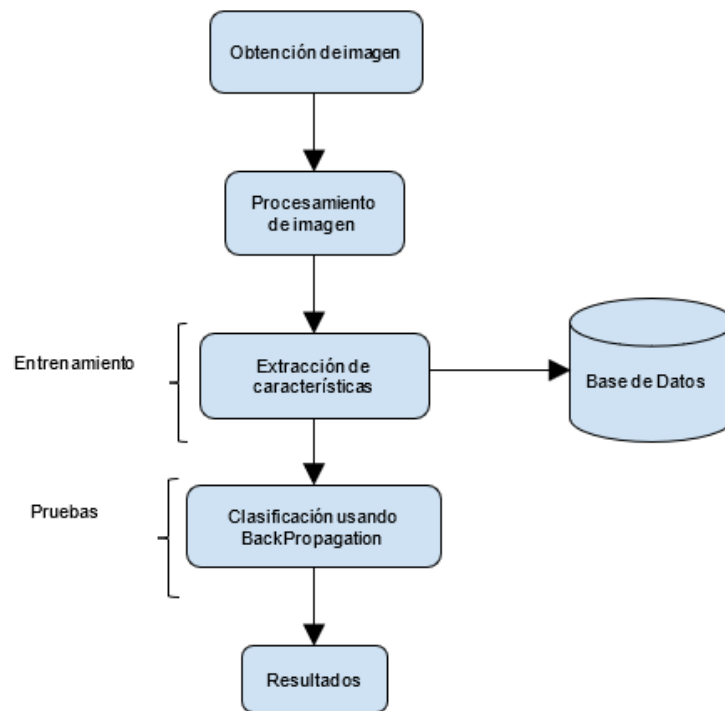


Fig. 1. Esquema de clasificación de gestualidades.

La extracción de la característica se hace usando la transformación 2D discreta del coseno (DCT 2D), ya que este método permite representar cualquier señal en el plano de las frecuencias [1]. Aplicando mencionada transformación en este trabajo se logró cambiar los datos de una imagen del dominio espacial al dominio de la frecuencia. Luego, de esta manera se formó un vector de características para entrenar la red de reconocimiento de patrones, es decir, la red neuronal Feedforward.

El clasificador es la red neural de Feedforward que utiliza el algoritmo de BackPropagation para entrenar al clasificador para datos de entrada contra datos de destino ya especificados, es decir, las cuatro emociones. La función de entrenamiento utilizada para esta red fue Trainscg, que es una función escalada de gradiente conjugado [1].

Para un clasificador, los valores de datos de entrada son características extraídas que forman un vector de particularidades y los datos de destino son otro vector 4X1. El parámetro T es una matriz para los datos objetivos donde las emociones ira, alegría, tristeza y susto se denotan por primera, segunda, tercera y cuarta filas de matriz respectivamente. Para una imagen que contenga una emoción de susto, este vector T mantendrá valores de 0 en todas las filas excepto 1 en la cuarta fila. A continuación se muestra el detalle del vector T.

$$T = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (1)$$

De esta manera el clasificador es entrenado para múltiples imágenes permitiendo al sistema aprender varias emociones.

IV.RESULTADOS

En la figura 2, se presenta la ventana principal de la aplicación. Esta ventana destaca todos y cada uno de los controles que permiten la interacción con la base de datos donde reposan las imágenes.



Fig. 2. Vista inicial de la aplicación.

Una vez conocidos los controles de la aplicación se podrá cargar una imagen al panel de entrada desde la base de datos previamente almacenada para su posterior análisis (Fig. 3).



Fig. 3. Obtener imagen desde la base de datos.

Previa obtención de la imagen, luego se procede a entrenar la red neuronal. Se debe elegir el tipo de emoción que presenta la imagen precargada. Cabe señalar que cada imagen es usada una sola vez para evitar el sobre-entrenamiento de la red (Fig. 4).



Fig. 4. Entrenamiento de la red neuronal.

Por último, para comprobar que el entrenamiento de la red ha sido el correcto se debe volver a cargar una imagen de prueba para que la aplicación determine qué tipo de emoción presenta el individuo. La imagen de prueba puede ser cualquiera elegida de forma aleatoria de entre un grupo de imágenes que correspondan al individuo en estudio (Fig. 5).



Fig. 5. Reconocimiento de emoción.

V.CONCLUSIONES

1.La gestualidad de las personas es de fácil reconocimiento siempre que la individualidad de la misma lo permita, ya que existen personas de difícil expresividad que no manifiestan sus estados de ánimo, en cuyos casos resulta de mayor dificultad.

2.La gestualidad de las personas es una característica inevitable, que permite caracterizar las formas de sus emociones y las manifestaciones de las mismas.

3.Las redes neuronales artificiales permiten resolver problemas de no linealidad, como el caso de la gestualidad humana, que difiere considerablemente en cada individuo.

4.Es preciso señalar que debido a que la muestra de estudio es reducida no es posible realizar un análisis mucho más profundo. Este punto sería una de las limitaciones de este trabajo.

REFERENCIAS

- [1] S. Gangwar, S. Shukla, D. Arora. "Human Emotion Recognition by Using Pattern Recognition Network", *Journal of Engineering Research and Applications*, Vol. 3, Issue 5, pp.535-539, 2013.
- [2] K. Rohit. "Back Propagation Neural Network based Emotion Recognition System". *International Journal of Engineering Trends and Technology (IJETT)*, Vol. 22, N° 4, 2015.
- [3] S. Eishu, K. Ranju, S. Malika, "Speech Emotion Recognition using BFO and BPNN", *International Journal of Advances in Science and Technology (IJAST)*, ISSN2348-5426, Vol. 2 Issue 3, 2014.
- [4] A. Fiszlelew, R. García-Martínez and T. de Buenos Aires. "Generación automática de redes neuronales con ajuste de parámetros basado en algoritmos genéticos". *Revista del Instituto Tecnológico de Buenos Aires*, 26, 76-101, 2002.
- [5] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel. "Handwritten digit recognition with a back-propagation network". In *Advances in neural information processing systems*. pp. 396-404, 1990.
- [6] G. Bebis and M. Georgiopoulos. "Feed-forward neural networks". *IEEE Potentials*, 13(4), 27-31, 1994.
- [7] G. Huang, Q. Zhu and C. Siew. "Extreme learning machine: a new learning scheme of feedforward neural networks". In *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference*. Vol. 2, pp. 985-990. IEEE, 2004.
- [8] D. Montana and L. Davis. "Training Feedforward Neural Networks Using Genetic Algorithms". In *IJCAI*, Vol. 89, pp. 762-767, 1989.
- [9] I. Sutskever, O. Vinyals and Q. Le. "Sequence to sequence learning with neural networks". In *Advances in neural information processing systems*. pp. 3104-3112, 2014.
- [10] J. Schmidhuber. "Deep learning in neural networks: An overview". *Neural networks*, 61, 85-117, 2015.
- [11] R. Santos, M. Rupp, S. Bonzi and A. Filetia, "Comparación entre redes neuronales feedforward de múltiples capas y una red de función radial para detectar y localizar fugas en tuberías que transportan gas". *Chem. Ing. Trans* , 32 (1375), e1380, 2013.