VOL. 16, NO. 18, SEPTEMBER 2021 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

INTELLIGENT SYSTEM FOR THE RECOGNITION OF FACIAL EMOTIONS: A TOOL FOR PEOPLE WITH AUTISM SPECTRUM DISORDER

Ferley Medina Rojas[1], Juan A. Castro Silva[1] and Faiber Robayo Betancourt[2]
[1]Departamento de Ingeniería de Software, Facultad de Ingeniería, Universidad Surcolombiana, Neiva, Huila, Colombia
[2]Departamento de Ingeniería Electrónica, Facultad de Ingeniería, Universidad Surcolombiana, Neiva, Huila, Colombia
E-Mail: ferley.medina@usco.edu.co

ABSTRACT

Autism spectrum disorder (ASD) is a neuropsychological disorder that begins at an early age in people and prevails throughout life. It affects the ability to communicate and interact with other people; they do not interpret facial emotions (Scared, Disgusted, Happy, Sad, Angry, Surprised, Contempt, and Neutral) like other people, resulting in isolation. Emotions help communicate to understand what other people feel. This paper presents a model for the recognition of emotions through facial expressions. It makes use of convolutional neural networks based on the VGG-Face architecture, trained to recognize faces and predict 8 emotions. The dataset used contains 13,000 images of people around the world. The precision obtained from the model was 84%, a value that will help people with ASD to recognize emotions so that they can interact in society.

Keywords: autism spectrum, recognition facial, convolutional neural networks.

1. INTRODUCTION

ASD is a developmental and neurological condition that begins in childhood and is prevalent throughout life. It affects how a person behaves, interacts with others, communicates, and learns [1]. The dissociative pattern in the face of interactions in a social environment is the main characteristic of people with ASD. During social development, the person is isolated and without the ability to interact with other people, much of this social isolation is due to the lack of communication and understanding of social processes, within these social activities that people with ASD are recognition of emotions because they lack emotional awareness which makes social and communicative interaction with other people even more difficult [2].

Facial recognition has a wide range of applications such as authenticating people and recognizing emotions through facial expressions. Currently, there are several companies specialized in the recognition of facial emotions such as Microsoft's Oxford project [3] is a catalog of API (Application Programming Interface) of artificial intelligence focused on computer vision, this system only works with photos, it is capable of to recognize seven central and neutral emotions (anger, disgust, fear, joy, sadness, and surprise). The [4] Cloud Vision API has the ability to detect faces along with facial attributes and emotional states present in an image.

Also, the [5] Amazon Rekognition API can detect, analyze, and compare faces for a wide variety of use cases, such as user verification, cataloging, people counting, and public safety. It also detects emotions such as happiness, sadness, or surprise. Through the use of computer vision, a “smart mirror” system was developed to identify basic and universal emotions (scared, disgusted, happy, sad, angry, surprise, contempt) as their absence. He uses face images labeled according to their facial expression and neural networks [6].

In this context, the project "Intelligent system for the recognition of facial emotions: a tool for people with autism spectrum disorder" is presented, which seeks to help people with a degree of difficulty in interpreting emotions so that they can have a normal social interaction. The identification of emotions for the understanding of what others feel guides to correct behavior and social formation [7]. With the cell phone camera, the image is captured so that the model developed and implemented in the APP interprets the information and presents it to the individual with ASD to facilitate communication with other members in real-time with an accuracy of 84%.

2. MATERIALS AND METHODS

2.1 Pre-Processing

The information is extracted from the Person the world data set that has 13719 images of different famous people in the world, with 8 emotions (Scared, Disgusted, Happy, Sad, Angry, Surprised, Contempt and Neutral). The images are divided into three groups: training with 85%, validation with 10% and test with 5% as shown in Table-1.

Table-1. Dataset conformation.

Categories	Percentage	Amount
Training	85%	11661
Validation	10%	1371
Test	5%	687

However, some of the categories of the data set contain few images such as: Fear and contempt. The Image Data Generator class implemented in the Keras library [8] was used to increase the number of training images, resizing and highlighting the images.

VOL. 16, NO. 18, SEPTEMBER 2021 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

As 50% of the image represents useful information, the problem was solved using the Haar Feature-based Cascade Classifiers algorithm [9] incorporated in the Python Keras pre-prossecing library. The images are resized [10] [11], only the face contour is captured, and the background is removed as shown in equation 1.

$$H(x,y) = \sum_p I\ (x,y) - \sum_n I\ (x,y) \qquad (1)$$

2.2 Feature extraction

It is done to extract the characteristics of the images in vectors that facilitate the classification. Task carried out with Convolutional Neural Networks (CNN) [12] implemented in Keras, with the Framework that uses the Google Tensorflow library as a backend.

Different feature extractors were designed all based on the VGG-Face network [13], due to the complexity of designing and training a CNN from scratch, this network has been trained to recognize people. However, the last layers, such as Fully Connected, were modified to extract specialized image descriptors as shown in Figure-1 using equation 2.

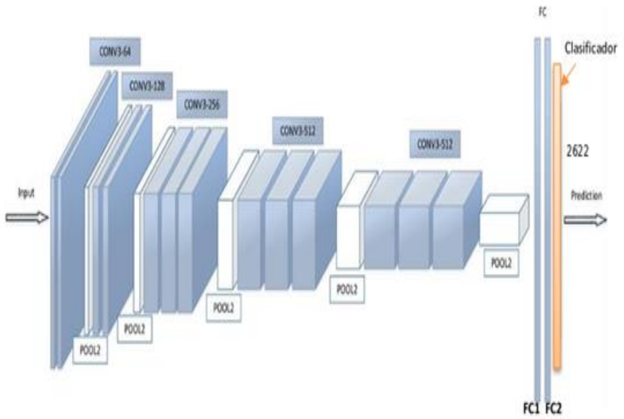

Figure-1. VGG-Face architecture

$$E(M') = \sum_{(a,p,n)\in T} max\{0, \alpha - |x_a - x_b|_2^2 + |x_a - x_p|_2^2\}, \qquad (2)$$

$$x_i = M' \frac{\phi(\ell_i)}{|\phi(\ell_i)|_2}$$

The last 4 layers of the VGG-Face network are used, made up of 3 convolutional layers and a max-pooling with an input of 224 x 224 and an output of 7x7. Also, 3 dense layers are added: one flatten, another with ReLu activation, and the last Dropout, obtaining an output of 1024 parameters using equation 3.

$$y_i = \sum_{j=1}^{n} E_i x_j - \theta_i \qquad (3)$$

2.3 Classifier

Once the characteristics of the images are extracted, the behavior of the system is evaluated with the Softmax classifier [14] implemented in the Python Keras library. The output parameters were classified into the 8 classes of emotions, with the probability of belonging of the images to each class using equation 4.

$$\sigma(y)_j = \frac{e^{y_j}}{\sum_{k=1}^{K} e^{y_j}} \qquad (4)$$

Adaptive moment estimation (ADAM) [15], a Backpropagation algorithm [16] that improves classifier learning, is used as optimizer. The method is easy to implement, computationally efficient, has low memory requirements, is invariable for diagonal scaling of gradients, and suitable for problems that are large in terms of data and / or parameters as described in equation 5.

$$y_t = y_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \varepsilon}} \qquad (5)$$

2.4 Evaluation

The algorithm is evaluated with the confusion matrix [17] that is implemented in the Python Keras Metrics library. This provides information on where classification errors occur and in which classes.

Accuracy: Refers to how close the result of a measurement is to the true value. In statistical terms, the accuracy is related to the bias of an estimate. It is also known as True Positive. It is represented by the proportion between the real positives predicted by the algorithm and all the positive cases; it is defined by equation 6.

$$acc = \frac{a+d}{a+b+c+d} \qquad (6)$$

3. RESULTS AND DISCUSSIONS

The model was trained in a paper space cloud computing platform in an environment with an Nvidia Quadro M4000 graphics card. The testing was carried out with 15% of the data. Three simulations were performed with network layer values from 1024 to 4096, varying the precision percentages from 81% to 84%, as shown in Table-2.

Table-2. Descripción de las simulaciones.

Simulation	Number of layers in the network	Accuracy
1	1024	81%
2	2048	82%
3	4096	84%

The highest precision obtained was 84% with 4096 layers of the neural network. For this result, two models are shown, the loss model Figure-2, from epoch

VOL. 16, NO. 18, SEPTEMBER 2021 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

three the deviation is almost constant with the training value, obtaining the minimum value in epoch 6. The precision model Figure-3 shows that from epoch 2 to 6 the deviation between training and observed values tends to be almost equal.

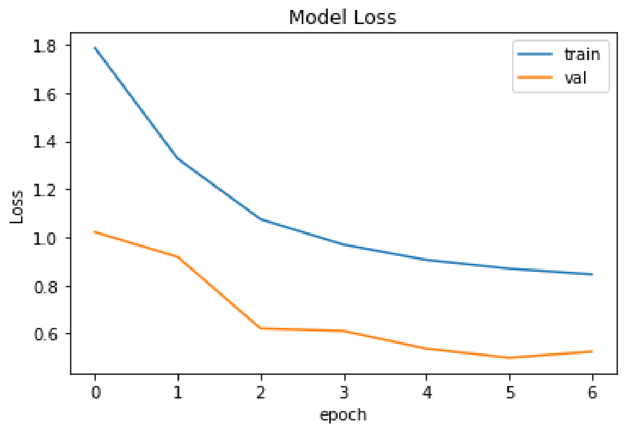

Figure-2. Loss model.

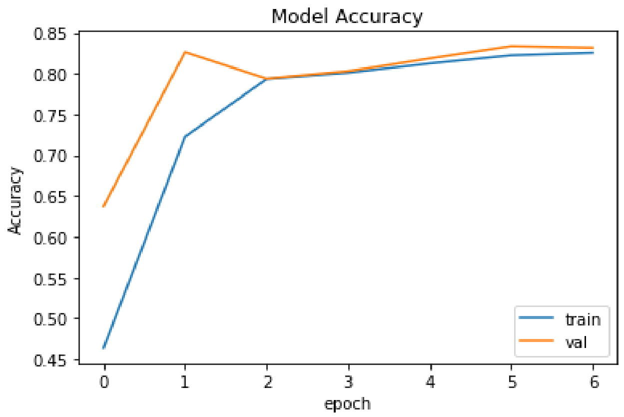

Figure-3. Precision model

The model was implemented in an APP composed of three buttons, one to capture the image, another to obtain the type of emotion, and the last one to exit. The APP was tested with three people who have ASD; all recognized the advantage of having the application available in the communication process. No further testing was possible due to people's unwillingness.

4. CONCLUSIONS

The implementation of the model in an APP with an accuracy of 84% in the prediction of emotions in people with ASD helps make the communication process easy, they can recognize eight types of emotions present in them, they improve their quality of life when interacting with an assertively way between them.

REFERENCES

[1] Asociación Americana de Psiquiatras. 2013. Guía de consulta de los criterios de diagnóstico del DSM 5, Arlinton Estados Unidos de América (EE. UU.).

[2] M. Miguel, A. María. 2006. El mundo de las emociones en los autistas. Revista Electrónica Teoría de la Educación: Educación y Cultura en la Sociedad de la Información. 7(2): 169-183.

[3] B. Smith and H. Shum. 2018. El Futuro Computarizado La Inteligencia Artificial y su rol en la sociedad, Washington Estados Unidos de América (EE. UU.), Microsoft Corporation Redmond.

[4] Google Cloud. 2021. API Cloud Vision Detect Faces, Link: https://cloud.google.com/vision/docs/detecting-faces.

[5] Amazon Rekognition. 2019. Guía del desarrollador, Link: https://docs.aws.amazon.com/es_es/rekognition/latest/dg/what-is.html.

[6] Gao Yao, JC. 2020. Diseño y desarrollo de un prototipo de espejo inteligente (SMIRROR). http://hdl.handle.net/10251/151906.

[7] Mora F., Sanguinetti A.M. 2004. Diccionario de neurociencia. Madrid. Alianza Editorial.

[8] V. Ivanchuk. 2018. Desarrollo de un modelo de reconocimiento de expresiones faciales mediante redes neuronales convolucionales y aprendizaje transferido. Trabajo de grado, Universidad Politécnica de Madrid, Madrid, España.

[9] A. Sharifara, M. S. Mohd Rahim and Y. Anisi. 2014. A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection. International Symposium on Biometrics and Security Technologies (ISBAST). Kuala Lumpur. pp. 73-78.

[10] FM Rojas, JA Castro, FR Betancourt. 2020. Intelligent system for the detection of iron stain on coffee growing leaves. ARPN Jornal of Engineering and Applied Sciences. 15(9).

[11] FM Rojas, JA Castro, FR Betancourt. 2020. Intelligent model for the detection of the phytophthora in the cocoa cropping black cob. ARPN Jornal of Engineering and Applied Sciences. 15(12).

ARPN Journal of Engineering and Applied Sciences

[12] O. Picazo Montoya. 2018. Redes neuronales convolucionales profundas para el reconocimiento de emociones en imágenes. M.S Tesis, Universidad Politécnica de Madrid, Madrid, España.

[13] Parkhi O.M., Vedaldi A. and Zisserman A. 2015. Deep Face Recognition. Proceedings of the British Machine Vision Conference (BMVC), Swansea.

[14] Wolfe J., Jin X., Bahr T. & Holzer N. 2017. Application of softmax regression and its validation for spectral-based land cover mapping. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences. 42.

[15] Kingma D. and Ba J. 2014. Adam: A method for stochastic optimization. Cornell University.

[16] Goodfellow Ian, Bengio Yoshua, Courville Aaaron. 2016. Deep Learning. MIT Press. p. 196.

[17] B. Giovanni. 2019. Evaluación de la eficiencia de los modelos machine learning para la predicción de la calidad del software desarrollado en IBM RPG usando la matriz de confusión y las curvas roc.