Journal of Engineering and Applied Sciences 12 (20): 5332-5337, 2017
ISSN: 1816-949X

Colorimeter Using Artificial Neural Networks

Laura Pramparo and Robinson Jimenez Moreno
Faculty of Engineering, Nueva Granada Military University, Bogota, Colombia

Abstract: The following study presents the development of a color classification algorithm for convolutional neural networks and fully-connected neural networks which uses a database of 200 images per color and between 12 and 18 colors to be classified for the training of the two networks. Subsequently, a comparison was made between their accuracy percentages where the best results were 95.33% for the convolutional neural network and 93.33% for the fully connected in the recognition of 12 colors and 93.67 and 35.23% for 18 colors, respectively. Finally, the best network is selected to design a video recognition application and the results are presented.

Key words: Convolutional neural network, fully-connected neural network, colorimeter, neural network architectures, colors

INTRODUCTION

A Fully-Connected Neural Network (FCNN) is a supervised learning algorithm whose structure seeks to mimic the behavior of human neural networks by adjusting scalar variables, known as weights that define the intensity of connection between one neuron and another as explained in detail by Riedmiller and Braun (1993) and Demuth *et al*. (2014). As explained by Nose-Filho *et al*. (2011) and Liu *et al*. (2017), the network is composed of multiple layers, each of which has a specific number of neurons that connect to all the neurons of the adjacent layers and are trained by reading an array of input and output variables and the readjustment of the value of each of the network weights.

The FCNN as well as the convolutional neural networks which are summarized below have been used for image recognition applications as by Rowley *et al*. (1998), Stathopoulou and Tsihrintzis (2004) where the presence of a face in a grayscale image is sought by means of a sliding window that performs a 20×20 pixel crop to enter the network and try to extract some of its main features like the eyes, nose and mouth and indicate whether or not there is a face. On the other hand by Golomb *et al*. (1991) in addition to doing face recognition, it is tried to classify each according to their gender where they manage to train a network with a recognition error of 8.1% and by Ming-Jung *et al*. (2003) and Phung *et al*. (2005) trained a FCNN to detect and classify different skin tones in an image.

On the other hand, Convolutional Neural Networks (CNN) (Zeiler and Fergus, 2014) are supervised learning algorithms that specialize in the recognition of patterns in images such as in the recognition of handwritten characters (Simard *et al*., 2003) or even in the recognition of 1000 different categories of images (Krizhevsky *et al*., 2012; Simonyan and Zisserman, 2014). Another recent application of them are focused in speech recognition (Abdel-Hamid *et al*., 2014), even developing very deep CNN for the recognition of speech in different languages as shown by Sercu *et al*. (2016). Unlike the FCNN, a CNN does not need the interconnection of large numbers of neurons, instead it uses filters of varying sizes to analyze an image by means of convolutions. However, although, this one has great capacity to process the different patterns of an image, this network has not been implemented to act as a colorimeter, although, some applications that have an approach to concepts of convolution have been developed as discussed by Cardei (2000) where they use the Retinex algorithm to solve camera color consistency problems through neural networks. Considering this, one can think that applying it to the recognition of colors could be useful, even achieving an efficient behavior in the recognition of patterns of different colors that have similarities to each other which will be developed in the present work.

In this study, two types of artificial neural networks (FCNN and CNN) are trained for the recognition of 12 and 18 colors and the results are compared between them to select the most efficient network and to use it in the recognition of color in objects captured by video.

MATERIALS AND METHODS

Colormeter database: The algorithm developed enables the training of a convolutional neural network and a fully-connected neural network for color recognition, both for plain color images and for colored objects on a white background. For the training of these two networks, two

Corresponding Author: Laura Pramparo, Faculty of Engineering, Nueva Granada Military University, Bogota, Colombia

Table 1: Coding of patterns for classification of 12 colors

Colors	Codes
Light yellow	00000
Intense yellow	00001
Blue	00010
Dark blue	00011
Lilac	00100
Purple	00101
Orange	00110
Black	00111
Red	01000
Light pink	01001
Intense pink	01010
Red wine	01011

Table 2: Coding of patterns for classification of 18 colors

Colors	Codes
Light yellow	00000
Intense yellow	00001
Blue	00010
Dark blue	00011
Dark brown	00100
Gray	00101
Lilac	00110
Purple	00111
Orange	01000
Black	01001
Red	01010
Light pink	01011
Intense pink	01100
Sand	01101
Lime green	01110
Sea green	01111
Dark green	10000
Red wine	10001

databases were formed, one for the classification of 12 colors composed of the colors of Table 1 and another for the classification of 18 colors composed of the colors of Table 2, obtaining a total of 2400 and 3600 images, respectively. Each image contains a sample of the color in different elements such as T-shirts, shoes, plain colors, etc. with a uniform white background to avoid being confused with other colors, since the purpose is to recognize a single color per sample. To standardize the database, each image is scaled to a size of 64×64 pixels, since the input of the networks is also of that size.

To validate the trained networks, a validation database of 50 images per category is also built. A sample of the training images is illustrated in Fig. 1. The networks will be trained and tested and finally the network with the best percentage of accuracy and the ability to classify 18 colors is selected for a color recognition application by means of a video capture.

Fully-connected neural network: As observed, a FCNN presents a complete connection between all neurons of one layer with the next which means that each pixel of the input image is connected to each neuron of the first layer of the network, i.e., the 4096 pixels of the input image (64×64 pixels) are each connected to all the neurons of the first layer which considerably increases the computational requirement required to train only one layer with 4096×N

Fig. 1: Training images

weight where N is the number of neurons of said layer. Consequently, the increase of layers in the network would raise in a great amount the weights to calculate, reason why the training would become too long and delayed.

In response to this drawback, the training of the network was focused on the main problem of the application which consists of the classification of the image according to its color. Therefore, an image preprocessing phase was added where the main color of the input image is extracted by reading each of its pixels and calculating the average of all those pixels whose color is different from the target.

With Eq. 1 it was calculated the average of the main color of each image where im is the input image, N the number of pixels with color, RGB the average of each color component, f the number of rows in the image, I, j, n the number of columns in the image are the iterators where varies from 1-3 to select the RGB component.

$$\mathrm{RGB}(n) = \frac{\sum_{i=0}^{f}\sum_{j=1}^{c} \mathrm{im}(i,j,n)}{N} \quad (1)$$

With the color averages of each image, an X array of 3 rows per M columns was mounted where M is the number of training images and entered into the network as input patterns. Each column of the array contains the average color of each input image and the 3 rows correspond to the RGB components of that color. Each pattern was assigned a binary code to classify it as shown in Table 1 and 2 and a Y array of 5 rows per M columns was created where 5 equals the number of bits in the binary code.

Because the hyperbolic tangent was used as the activation function from the training binary code was all zeros were changed by -1 to cover the entire working range of the activation function. The hyperbolic tangent allows to handle a wide range of work for the training of the network, since its outputs range from -1 to 1 while other activation functions like the sigmoid have a short range from 0-1, making the outputs very close to each other.

Table 3: Percentages of accuracy for classification of 12 and 18 colors by means of an FCNN

No. of colors	Accuracy (%)
12	93.33
18	35.23

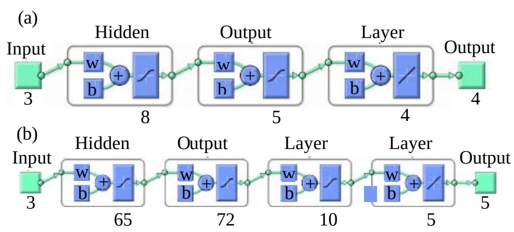

Fig. 2: Architecture of the multilayer neural network with backpropagation to classify: a) 12 colors and b) 18 colors

The X and Y arrays were used as input and output patterns, respectively, in network training and the MATLAB® Toolbox was used to calculate all weights. There were two different trainings, one to recognize 12 colors (Network 1) and another to recognize 18 colors (Network 2) where the network architecture of Fig. 2a was used to train the 12 colors and the network of Fig. 2b to train the 18 colors.

Table 3 shows the accuracy percentages obtained for each network and Fig. 3 shows the recognition of both networks where the colors of the top of Fig. 3a and b are the training colors and the bottom is divided into test and classification colors of the network 1 and 2, respectively. The classification colors are those with which the network classified each of the test colors using the previously trained weights.

As can be seen, network 1 managed to recognize almost all test colors, achieving 93% accuracy but with increasing the number of patterns to 18, accuracy was reduced to 35%, confusing almost all colors and unable to recognize almost any of the green tones. Network 2 generalized the colors yellow, blue and purple and the others did not manage to classify them. The number of neurons per layer was varied to try to improve the recognition but the accuracy percentages were around 22%.

Convolutional neural network: To do the recognition of the colors, a CNN architecture was implemented as shown in Fig. 4 which does not need to be very deep to be able to recognize the colors, since as such it does not require to obtain so detailed patterns but different shades of the same color that when varying, it can resemble another color that is in the categories. For the recognition of 12 and 18 categories, the same architecture is used.

Once the network has been trained, the validation is performed, obtaining the results presented in Table 4 where the accuracy percentages for the trained network

Table 4: Percentages of accuracy for 12 and 18-color classification by CNN

No. of colors	Accuracy (%)
12	95.33
18	93.67

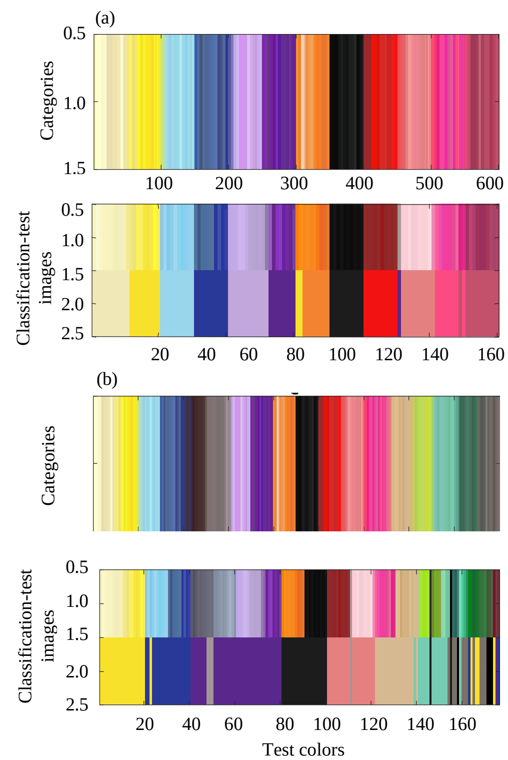

Fig. 3: Classification for test colors: a) 12 colors and b) 18 colors

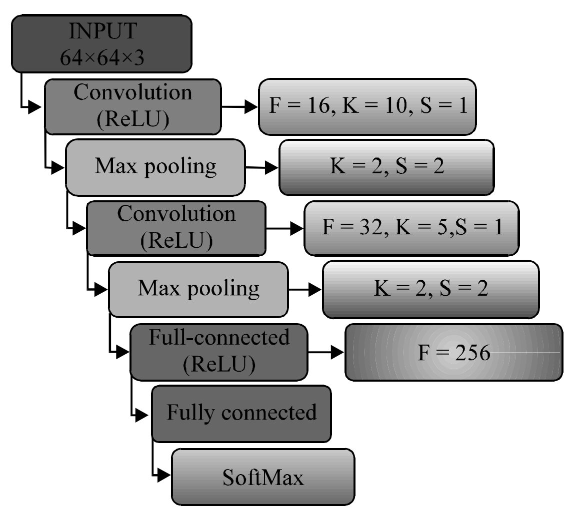

Fig. 4: CCN architecture used for the colorimeter where F represents the number of filters, K is the kernel or size of the filters and S is the step with which the filter moves

with 12 colors as for 18 colors are found. Taking into account these results, it can be seen that the architecture trained for both 12 and 18 colors maintains an accuracy above 93% which is a very good accuracy as to the application in which it is to be used.

The main difficulty to reach 100% are the similarities in the tonalities of some colors as for example the color Sand which tends to be confused with the pale yellow or the black with the colors cataloged as "dark", since these contain parts of shades close to black due to the shadows contained in the images of these colors.

RESULTS AND DISCUSSION

Analysis of validation results: The convolutional neural network managed to obtain a greater percentage of accuracy in the recognition of colors against the FCNN, both to recognize 12 colors and to recognize 18. In addition, the convolutional neural network does not require a preprocessing phase to extract the colors of each image before the training, since it itself trains each of its filters to extract the colors of the images and classify them, surpassing external agents like shades or secondary colors which in FCNN, change the average color of the images that enter to them and affect the result of classification.

The convolutional neural network, on the other hand is designed to read and process images, using convolutional filters that help reduce the computational requirement presented in the FCNN which makes it an ideal network for image color classification. However, it is important to have training images that allow the network to generalize its classification, i.e., to actually learn to classify by color and not by the shape of objects or other characteristics that may affect training which is achieved by the use of images with different forms, dimensions, rotations and positions but with a single characteristic in common that is the color. Such complications do not occur in the FCNN since it only looks at the tonalities of each pixel, not the image in its hue, so it is ensured that it learns what the user wants.

Video tests: Taking into account the analysis carried out previously, the convolutional neural network was chosen for the realization of the tests in real time, since, although, the number of categories was increased, it obtained an accuracy response higher than that obtained with the Fully-connected network.

For the tests, a controlled environment with a uniform white background and a fixed illumination was created, in order to avoid de variation of the parameters of the webcam used (brightness, red intensity, exposure, etc.) and thus avoid changes in the tonalities of the objects to which the color will be recognized. Figure 5 shows the tests performed on different objects in real time, however, although, the environment was controlled, on several occasions the webcam used auto-adjusted its parameters, even changing the tone of the background, as happened with the color orange where because of the intensity of the color, the webcam changed the saturation parameter changing the color of the environment captured. Despite this the neural network responded adequately and recognized all the colors satisfactorily, even differentiating the 2 closet tonalities that were light yellow (Bright-Yellow) and intense yellow (Yellow).

Fig. 5: Real-time tests for object color recognition

CONCLUSION

A comparison was made between two different types of networks executed for the same purpose, the recognition of a certain quantity of colors where it was shown that the CNN has better characteristics to be implemented in said application, since it does not need an initial image preprocessing as detailed as required by the FCNN which in turn reduces the computational cost required, since the FCNN handles a large number of connections between layers, since all of its neurons are connected.

On the other hand by adding more colors, the CNN accuracy response was higher than the FCNN with a difference of 2% for the classification of 12 colors and 58.44% for 18 colors that demonstrates that among a greater number of colors, the CNNs present a better behavior than the FCNNs.

The main fault in color recognition, according to the tests performed is when there are different changes of light which causes changes of color tonality in the objects, making them look like other training colors. Another influential factor is the quality of the camera used which in this case was a webcam which when detecting strong colors, showed changes of brightness and color nuances, making even the background will change its tone.

Additionally, for the selection of the images of the database, a variable that cannot be controlled which is the perception of the color of the people, must be faced what makes that a single color is described or taken with different perspectives, for instance, the lime green color which according to the person, can vary from a green with strong tones of yellow to a light green, resembling that of a lemon.

The application of specific color recognition had not been implemented previously through neural networks which allows to extend the variety of applications in which the neural networks can be used, mainly the CNN which can be a great utility in several future works.

REFERENCES

Abdel-Hamid, O., A.R. Mohamed, H. Jiang, L. Deng and G. Penn *et al.*, 2014. Convolutional neural networks for speech recognition. IEEE. ACM. Trans. Audio Speech lang. Process., 22: 1533-1545.

Cardei, V.C., 2000. A neural network approach to colour constancy. Ph.D Thesis, Simon Fraser University, Burnaby, British Columbia.

Demuth, H.B., M.H. Beale, D.O. Jess and M.T. Hagan, 2014. Neural Network Design. 2nd Edn., Martin Hagan, USA., ISBN:0971732116,.

Golomb, B.A., D.T. Lawrence and T.J. Sejnowski, 1991. Sexnet: A neural network identifies sex from human face. Adv. Neural Inform. Process. Syst., 3: 572-577.

Krizhevsky, A., I. Sutskever and G.E. Hinton, 2012. Imagenet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems, Leen, T.K., G.D. Thomas and T. Volker (Eds.). MIT Press, Cambridge, Massachusetts, USA., ISBN:0-262-12241-3, pp: 1097-1105.

Liu, W., Z. Wang, X. Liu, N. Zeng and Y. Liu *et al.*, 2017. A survey of deep neural network architectures and their applications. Neurocomputing, 234: 11-26.

Ming-Jung, S., V. Deepthi and V.K. Asari, 2003. Neural network based skin color model for face detection. Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop, Oct. 15-17, IEEE Computer Society, USA., pp: 141-145.

Nose-Filho, K., A.D.P. Lotufo and C.R. Minussi, 2011. Preprocessing data for short-term load forecasting with a general regression neural network and a moving average filter. Proceedings of the 2011 IEEE International Conference on PowerTech Trondheim, June 19-23, 2011, IEEE, Trondheim, Norway, ISBN: 978-1-4244-8419-5, pp: 1-7.

Phung, S.L., A. Bouzerdoum and D. Chai, 2005. Skin segmentation using color pixel classification: analysis and comparison. IEEE. Trans. Pattern Anal. Mach. Intell., 27: 148-154.

Riedmiller, M. and H. Braun, 1993. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. IEEE Int. Conf. Neural Networks, 1: 586-591.

Rowley, H.A., S. Baluja and T. Kanade, 1998. Neural network-based face detection. Pattern Anal. Mach. Intell., 20: 23-28.

Sercu, T., C. Puhrsch, B. Kingsbury and Y. LeCun, 2016. Very deep multilingual convolutional neural networks for LVCSR. Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), March 20-25, 2016, IEEE, Shanghai, China, ISBN:978-1-4799-9988-0, pp: 4955-4959.

Simard, P.Y., D. Steinkraus and J.C. Platt, 2003. Best practices for convolutional neural networks applied to visual document analysis. Proceedings of the of 7th International Conference on Document Analysis and Recognition ICDAR Vol. 3, August 3-6, 2003, IEEE, New York, USA., pp: 958-962.

Simonyan, K. and A. Zisserman, 2014. Very deep convolutional networks for large-scale image recognition. Master Thesis, Cornell University, Ithaca, New York.

Stathopoulou, I.O. and G.A. Tsihrintzis, 2004. An improved neural-network-based face detection and facial expression classification system. Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics Vol. 1, October 10-13, 2004, IEEE, Hague, Netherlands, ISBN:0-7803-8566-7, pp: 666-671.

Zeiler, M.D. and R. Fergus, 2014. Visualizing and understanding convolutional networks. Proceedings of the European Conference on Computer Vision, September 6-12, 2014, Springer, Zurich, Switzerland, pp: 818-833.