Color Face Recognition by Using Quaternion and Deep Neural Networks

Abdelmajid EL ALAMI
Sidi Mohamed Ben Abdellah University
Fez, Morocco
elalami78@gmail.com

Zouhir LAKHILI
Sidi Mohamed Ben Abdellah University
Fez, Morocco
lakhili.zouhir@gmail.com

Abderrahim MESBAH
Sidi Mohamed Ben Abdellah University
Fez, Morocco
abderrahimmesbah@yahoo.fr

Aissam BERRAHOU
Mohammed V University
Rabat, Morocco
aissam.berrahou@gmail.com

Hassan QJIDAA
Sidi Mohamed Ben Abdellah University
Fez, Morocco
qjidah@yahoo.fr

***Abstract*— In this paper we propose a new model for color face recognition based on quaternion number and deep neural networks (DNN), to enhance the classification accuracy of color face recognition. The proposed model is derived by introducing image represented in quaternion domain as an input vector in DNN model, frequently, utilized in many vision tasks of pattern recognition. The utilization of quaternion representation (QR) aims to reduce the size of input vector and consequently the computational complexity. On the other hand, color information from the image while obtaining best classification results. The simulations results are performed on color face databases to demonstrate the effectiveness of the proposed model. The obtained results show that they outperform other existing algorithms.**

Keywords— quaternion, color face recognition, complexity, deep neural networks

I. Introduction

Object classification has received considerable attention in the field of computer vision, its concepts have been applied in many applications such as robotic vision, surveillance system, manufacturing, face recognition, etc. With recent advances in technologies of sensing hardware digital camera, almost all the objects captured are chromatic. However, a color image has the ability to transmit more information than a binary or monochrome image. The components of three colors red, green and blue (RGB) specified by a color space, are stored to each pixel of image, the color can be defined in various manners which conduct meaningfully different abilities in terms of discriminating power for classification task [1].

The classical approach for color images analysis is to process each channel of the color separately using gray scale technique, and to combine the three individual outputs to give the global results. Consequently, this approach not succeeds to capture the inherent correlation between three color channels, which give non optimal representation of color image [2-4].

Recently, quaternion algebra has been used in color image processing and recognition to represent pixels of color images in three imaginary parts of quaternion numbers which represent the red, green and blue [2, 5-10]. The benefit of the representation in the quaternion domain is that the color image can be dealt overall as a vector field [2, 5-7]. The quaternion algebra has been introduced in the field of color image processing by Sangwine [5], and Pei and Cheng [11], and has been used in multiple research area such as molecular modeling [12], segmentation [2] and image texture [13].

In this paper we propose a new model for color face recognition, which take features represented by quaternions from image color as an input vector in deep neural networks. In fact, the representation in the quaternion domain preserves pertinent color information from the image with high discrimination power. Hence, the introduction of the quaternion numbers in the DNN model makes it possible to design the model that reduces considerably the computational cost by decreasing the number of layers and parameters and to gives best classification scores. The proposed model is evaluated on four real datasets: faces94, faces95, faces96, and grimace. The experimental results indicate that our proposed model outperform others existing methods in terms of classification results.

The rest of this paper is organized as follow: section II an overview of some basic features of quaternion and quaternion color representation are presented. The details of the proposed model are explained in section III. Section IV gives the experimental results of the proposed model and the comparisons to others existing methods. Section V concludes the present work.

II. Quaternion Algebra

Quaternions, invented by the mathematician Sir William Rowan Hamilton [14] in 1843, are generalizations of complex numbers. A quaternion defined by one real part and three imaginary parts is presented as follows:

$$q = a + bi + cj + dk \tag{1}$$

Where $a, b, c, d \in R$, and i, j and k are three imaginary units that obeys the following rules:

$$\begin{aligned} &i^2 = j^2 = k^2 = -1,\ ij = -ji = k,\ jk = -kj = i, \\ &ki = -ik = j. \end{aligned} \tag{2}$$

If the real part $a = 0$, q is called a pure quaternion.

The addition and subtraction of tow quaternion $q_1, q_2 \in H$ are defined by:

$$q_1 \pm q_2 = (a_0 \pm b_0) + (a_1 \pm b_1)i + (a_2 \pm b_2)j + (a_3 \pm b_3)k \tag{3}$$

Where:

$q_1 = a_0 + a_1 i + a_2 j + a_3 k$ and $q_2 = b_0 + b_1 i + b_2 j + b_3 k$

The multiplication of tow quaternion $q_1, q_2 \in H$ is defined by:

$$\begin{aligned} q_1 \cdot q_2 = & (a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3) \\ & + (a_0 b_1 + a_1 b_0 + a_2 b_3 - a_3 b_2) i \\ & + (a_0 b_2 + a_2 b_0 + a_3 b_1 - a_1 b_3) j \\ & + (a_0 b_3 + a_3 b_0 + a_1 b_2 - a_2 b_1) k \end{aligned} \tag{4}$$

The conjugate and modulus of a quaternion are respectively defined by:

$$\bar{q} = a - bi - cj - dk \tag{5}$$

$$\|q\| = \sqrt{a^2 + b^2 + c^2 + d^2} \tag{6}$$

A color image RGB, $f(x, y)$, can be represented using a pure quaternion, as follows:

$$f(x, y) = f_R(x, y) i + f_G(x, y) j + f_B(x, y) k \tag{7}$$

Where $f_R(x, y)$, $f_G(x, y)$ and $f_B(x, y)$ represent the red, green, and blue channel components, respectively, based on the definition of the quaternion.

The unit pure quaternion μ. Basically, μ can be defined as a linear combination of i, j and k such as:

$$\mu = \alpha i + \beta j + \gamma k, \alpha, \beta, \gamma \in R, \|\mu\| = 1 \tag{8}$$

μ has the value $\mu = \frac{1}{\sqrt{3}}(i + j + k)$ corresponds to the grey line in RGB space, the three imaginary parts are weighted with the same factor $\frac{1}{\sqrt{3}}$ in order to provide the same privileges to the three color space channels.

By using Eq. (7) and $\mu = \frac{1}{\sqrt{3}}(i + j + k)$, the intensity function $F(x, y)$ of color images in right-side form is represented as:

$$F(x, y) = f(x, y)\mu \tag{9}$$

III. Proposed Model

In this section, we present a description of our model which takes images represented by quaternion as input vector. We give an overview of some functions which are used to improve the accuracy such as function of activation Relu, Elu and output function Softmax. The proposed model is depicted in Fig. 1 and detailed in Table I.

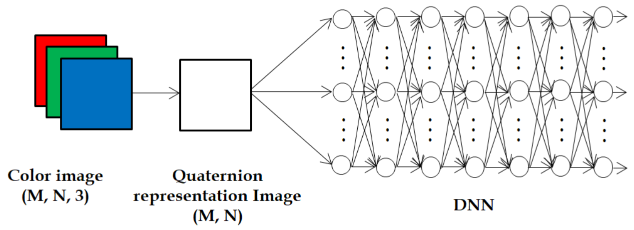

Fig. 1. Proposed DNN model

TABLE I. PROPOSED MODEL

Layer	Purpose	Activation
H0	Input image quaternion vector	M×N
H1	Fully connected+BN+Elu+Dropout	100
H2	Fully connected+BN+Relu+Dropout	170
H3	Fully connected+BN+Relu+Dropout	240
H4	Fully connected+BN+Relu+Dropout	170
H5	Fully connected+BN+Relu+Dropout	120
O	Softmax	Number of subjects

The model contains seven layers: an input layer, five hidden layers, and an output layer. The input vector of deep neural networks model has the size M×N. Five hidden layers with number of neurons 100, 170, 240, 170 and 120 are respectively used from H1 to H5 layer. The last softmax layer outputs correspond to the number of class labels. The proposed model contains three types of layers as follows:

- Input layer H0: the input layer of the proposed model is represented by quaternion number which is used to form the vector of every color image. the input vector of the proposed model can be written as:

$$V = F(x, y) / 0 \le x \le M - 1, 0 \le y \le N - 1 \tag{10}$$

- Hidden layer: Five Hidden layers from H1 to H5 contain successively 100, 170, 240, 170, 120 neurons. Hi can be denoted by the following formula:
- Hidden layer: Five Hidden layers from H1 to H5 contain successively 100, 170, 240, 170, 120 neurons. Hi can be denoted by the following formula:

$$Y_{Hi} = \Psi_{Hi}(b_{(Hi)} + W_{(Hi)} Y_{Hi-1}) \tag{11}$$

Where $Y_{H0} = V$, $b_{(Hi)}$ and $W_{(Hi)}$ are respectively the bias vector and the weight matrix of the i^{th} hidden layer Hi, and $\Psi_{Hi}(.)$ is an activation function used in Hi.

- ReLu: rectified linear unit is expressed mathematically as $f(x, y) = \max(0; x)$, it considerably accelerates the convergence of the stochastic gradient descent compared to other functions like Sigmoid or Tanh. Furthermore, Relu units can be fragile during training and avoids network saturation [15, 16].
- Elu: Exponential linear unit (ELU) differs from ReLu by negative inputs. For positive inputs, they are both in identity function form. On the other hand, Elu also have improved learning characteristics and avoid a vanishing gradient via the identity for positive values compared to the other activation functions [17].
- BN: Batch normalization can be applied after each layer. The goal is to: (1) accelerate training process, (2)

overcome the problem of overfitting and reduce the gradients dependencies [18].

- Dropout [19] was applied after each hidden layer in order to overcome the overfitting by regularizing deep networks. In this work a ratio of 0.15 (85% of the activations were kept), applied after each layer, in order to ameliorate the classification accuracy.

- Output layer: The output of the network provide the probabilities of the input corresponding to each element belonging to a label by using function softmax [20].

  The output function of our model is defined by the following formula:

$$g(V) = soft\max(b_{(H6)} + W_{(H6)} Y_{H5}) \tag{12}$$

IV. Experimental results

In this section, a set of experiments are conducted to demonstrate the classification performance of the proposed model. Four databases of color face recognition are chosen: faces94, faces95, faces96 and Grimace given by the University of Essex [21], the description of these databases will be presented in the next subsection. All experiments were performed in Google Colaboratory framework with Tesla K80 GPU.

A. Faces94 dataset

The face 94 [21] consists of 3040 color images of male and female with the size of 180×200 distributed over 152 classes in separate directories: 19 females, 113 males and 20 malestaff. Each class contains 20 images taken from different viewpoint and under various lightings. A few examples of images used are shown in Fig. 2. In the experiment each directory of datasets is divided in three splits. In the first we have selected randomly (50%) for training and the rest for test. In the second we have selected randomly (40%) for training set and (60%) for the test. In the third we have randomly selected (30%) for training and the remaining for testing.

Fig. 2. Some color images from faces94 dataset.

Table II presents the classification accuracy results on the three splits. It can be observed that the proposed model achieves the rate of classification of 100% for female and malestaff with all testing splits and more than 99.62% for male subset.

TABLE II. Classification results for three training splits on Faces94 dataset

Dataset		Accuracy		
		Test 50%	***Test 60%***	***Test 70%***
Faces94	Female 19 classes	100%	100%	100%
	Male 113 classes	99.91%	99.70%	99.62%
	Malestaff 20 classes	100%	100%	100%

B. Faces95 dataset

The faces95 datasets [21] consists of 1440 color images of male and female classified in 72 classes, each class contains 20 different color images with the size of 180×200. Fig. 3 shows some color images examples from the faces95 dataset. In the experiments, we have used 600 faces of 30 subjects to evaluate the performance of our proposed model. We have designated randomly 300 (50%) for training and the rest for testing.

Fig. 3. Some color images from faces95 dataset.

Table III presents the classification results for faces95 database and the comparison with other different algorithm in the state of art. It can be seen that the proposed model achieves better results with classification rate of 96.33%, as compared than other existing methods.

TABLE III. Comparison of classification accuracy of the proposed model with different methods on faces95 dataset.

Methods	Accuracy
ROTMIs+BPNN [24]	86.67%
QROTMIs+QBPNN [24]	91.67%
OFMMIs+BPNN [24]	91.00%
QOFMMIs+QBPNN [24]	93.33%
ZMIs+BPNN [24]	91.67%
QZMIs+QBPNN [24]	94.00%
PZMIs+BPNN [24]	92.67%
QPZMIs+QBPNN [24]	95.67%
QFMMIs [22]+BPNN [24]	90.00%
QBDPCA [23, 24]	86.33%
Our	96.33%

C. Faces96 dataset

The faces96 dataset [21] includes 3040 images color with the size of 196×196 distributed on 152 classes. Each subject contains 20 color images captured from various positions and under different lightings. A few face color images used are shown in Fig 4.

Fig. 4. Some face color images from faces96 datasets.

We have performed the classification on original dataset noted D1, which contains 3040 color images, and two other subset from faces96, the first noted D2: contains 600 images over 30 classes selected randomly from the original dataset. The second, noted D3, includes 2000 images on 100 individuals selected from the original faces96.

In the experiments, each dataset are divided in two splits: the first we have selected randomly 30% for the test set and 70% for the train set. In the second we have randomly selected 50% for train set and the remaining for test set.

Table IV depicts the classification results on three datasets D1, D2 and D3 respectively, it can be observed that the proposed model achieve the best classification results of 100% when using D2 datasets with 50% of training. Table V, and Table VI compare the performance of the proposed model to other methods on D2 and D3 respectively, it can be seen that our model with quaternion representation achieves the best classification results.

TABLE IV. CLASSIFICATION RESULTS ON THREE DATASETS USING OUR PROPOSED MODEL WITH TOW SPLITS.

Datasets	Accuracy	
	Training with 50%	*Training with 70%*
D1	99.07%	99.33%
D2	100%	100%
D3	99.69%	99.83%

TABLE V. COMPARISON OF CLASSIFICATION ACCURACY ON D2 DATASET WITH OTHER METHODS.

Methods	Accuracy
ROTMIs+BPNN [24]	94.00%
QROTMIs+QBPNN [24]	95.67%
OFMMIs+BPNN [24]	94.67%
QOFMMIs+QBPNN [24]	96.67%
ZMIs+BPNN [24]	96.33%
QZMIs+QBPNN [24]	98.33%
PZMIs+BPNN [24]	98.00%
QPZMIs+QBPNN [24]	99.00%
QFMMIs [22]+BPNN [24]	93.67%
QBDPCA [23, 24]	97.00%
Our	100 %

TABLE VI. COMPARISON OF THE CLASSIFICATION ACCURACY OF DIFFERENT METHODS ON D3 DATASET.

Methods	Accuracy
PFQDHM$_S$ [25]	93.92%
QDHM$_S$ [25]	94.61%
PFQBFM$_S$ [25]	83.34%
QBFM$_S$ [25]	82.5%
PFQZM$_S$ [25]	93.6%
QZM$_S$ [25]	90.6%
poFMD [25]	89.01%
LCDM [25]	74.74%
CM$_{50}$ [25]	64.34%
CM$_{10}$ [25]	55.58%
Our	99.83%

D. Grimace

The grimace dataset [21] is composed of 360 color images with the size of 180×200 distributed over 18 classes; each class includes 20 different color images. A few examples are depicted in Fig. 5. In the experiments we are randomly considered 180 (50%) objects for training set and the rest of objects for testing set.

Fig. 5. A few example of images from Grimace dataset.

Table VII shows the classification accuracy results for grimace dataset and the comparison to other methods. It can be seen that the proposed model achieves high result with classification rate of 100% as compared to some existing methods.

TABLE VII. COMPARISON OF RECOGNITION RATE OF THE PROPOSED MODEL WITH DIFFERENT METHODS ON GRIMACE DATABASE

Methods	Accuracy
ROTMIs+BPNN [24]	95.33%
QROTMIs+QBPNN [24]	97.44%
OFMMIs+BPNN [24]	96.67%
QOFMMIs+QBPNN [24]	99.44%
ZMIs+BPNN [24]	100%
QZMIs+QBPNN [24]	100%
PZMIs+BPNN [24]	100%
QPZMIs+QBPNN [24]	100%
QFMMIs [22]+BPNN [24]	98.33%
QBDPCA [23, 24]	99.44%
Our	100 %

E. Ablation and Time consuming

In this part, we present an ablation study on faces96. We compared the accuracy and time consuming in training step on faces96 dataset between our model with images of three channels (RGB) and the same model with the quaternion representation (QR) of image. Table VIII depicts the average of three computational times achieved after 1000 epochs. It can be observed that the use of quaternion representation with the proposed model is faster than the use of image on three channels (more than twice). Table IX shows slightly difference between these two methods in terms of classification rate.

TABLE VIII. COMPARISON OF TIME CONSUMING IN TRAINING STEP ON FACES 96 DATASET (WITH SPLIT 50% FOR TRAINING).

Method	Training time in second on three simulations			Average
	1	*2*	*3*	
With three channels (RGB)	621.48	620.60	684.93	642.33
With quaternion representation	266.15	264.75	266.37	265.75

TABLE IX. COMPARISON THE CLASSIFICATION RATE BETWEEN ABLATION AND QUATERNION REPRESENTATION ON FACES 96.

Datasets	Accuracy			
	Training with 50%		*Training with 70%*	
	Ablation	*With QR*	*Ablation*	*With QR*
D1	99.07%	99.07%	99.11%	99.33%
D2	100%	100%	100%	100%
D3	99.59%	99.69%	99.66%	99.83%

V. CONCLUSION

In this paper a new method for color face recognition was presented; it's based on DNN model by using images represented in quaternion domain as an input layer, this method was evaluated on real face databases, the simulated results demonstrate better performance and achieve high classification rates than other methods. Therefore, quaternion representations have better descriptive capabilities. As a future research, it will be interesting to use quaternion algebra in semantic segmentation and content-based image retrieval.

REFERENCES

[1] R. Lukac and K.N. Plataniotis, "Color image processing: methods and applications," CRC press, 2006.

[2] O.N. Subakan and B.C. Vemuri, "A quaternion framework for color image smoothing and segmentation," International Journal of Computer Vision. 91(3), pp. 233–250, 2011.

[3] A. Koschan and M. Abidi, "Digital Color Image Processing," John Wiley & Sons; 2008 Feb 15.

[4] Y. Xu, L. Yu, H. Xu, H. Zhang and T. Nguyen, "Vector sparse representation of color image using quaternion matrix analysis," IEEE Transactions on Image Processing 24(4), 2015, pp. 1315–1329.

[5] S.J., Sangwine, "Fourier transforms of colour images using quaternion or hypercomplex, numbers," Electronics letters 32(1), pp. 1979–1980, 1996.

[6] N.L. Bihan and S.J. Sangwine, "Color image decomposition using quaternion singular value decomposition," Visual Information Engineering, 2003. VIE 2003. International Conference on. IET, pp. 113–116, 2003.

[7] N.L. Bihan and S.J. Sangwine, "Quaternion principal component analysis of color images," In: Proc. 2003 10th IEEE Int. Conf. Image Processing (ICIP 2003), vol. 1, pp. 809–812, 2003.

[8] Y.F. Sun, S.Y. Chen and B.C. Yin, "Color face recognition based on quaternion matrix representation," Pattern Recognit. Lett. 32(4), pp. 597–605, 2011.

[9] B.J. Chen, H.Z. Shu, H. Zhang, G. Chen, C. Toumoulin, J.L. Dillenseger and L.M. Luo, "Quaternion zernike moments and their invariants for color image analysis and object recognition," Signal Process. 92(2), pp. 308–318, 2012.

[10] D. Rizo-Rodríguez, H. Méndez-Vázquez and E. García-Reyes, "Illumination invariant face recognition using quaternion-based correlation filters," J. Math. Imaging Vis. 45(2), pp. 164–175, 2013.

[11] S.C. Pei and C.M. Cheng, "A novel block truncation coding of color images by using quaternion-moment-preserving principle," In: Proc. IEEE Int. Symp. Circuits and Systems, vol. 2, pp. 684–687, 1996.

[12] C.F.F. Karney, "Quaternions in molecular modeling," Journal of Molecular Graphics and Modelling, vol. 25, no 5, pp. 595-604, 2007.

[13] D. Assefa, L. Mansinha, K.F. Tiampo, H. Rasmussen and K. Abdella, "Local quaternion fourier transform and color image texture analysis," Signal Process. 90(6), pp. 1825–1835, 2010.

[14] W.R. Hamilton, "Elements of Quaternions," Longmans Green, London, 1866.

[15] B. Xu, N. Wang, T. Chen and M. Li, "Empirical evaluation of rectifided activations in convolutional network," arXiv:1505.00853, 2015.

[16] X. Glorot, Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in 13th International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, (9), pp. 249-256, 2010.

[17] D. Clevert, T. Unterthiner and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUS)," in Conference ICLR2016, 2016.

[18] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in 32nd International Conference on Machine Learning, Lille, France, pp. 448-456, 2015.

[19] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," The Journal of Machine Learning Research, 15(1), pp. 1929-1958, 2014.

[20] P. Reverdy and N.E. Leonard, "Parameter Estimation in Softmax Decision-Making Models with Linear Objective Functions," IEEE Transactions on Automation Science and Engineering 13 (1): pp. 54–67. doi: 10.1109/TASE.2015.2499244

[21] L. Spacek, (2008, June 20). "Description of the Collection of Facial Images," [Online]. Available: http://cswww.essex.ac.uk/mv/allfaces/index.html

[22] L.Q. Guo and M. Zhu, "Quaternion fourier-Mellin moments for color image," Pattern Recognit. 44(2), pp. 187–195, 2011.

[23] Y.F. Sun, S.Y. Chen and B.C. Yin, "Color face recognition based on quaternion matrix representation," Pattern Recognit. Lett. 32(4), pp. 597–605, 2011.

[24] B. Chen, H. Shu, G. Coatrieux, G. Chen, X. Sun and J. L. Coatrieux, "Color image analysis by quaternion-type moments," Journal of mathematical imaging and vision, 51(1), pp. 124-144, 2015.

[25] N. Dad, N. En-nahnahi and S.E.A. Ouatik, "Parameter-free quaternary orthogonal moments for color image retrieval and recognition," Journal of Electronic Imaging, vol. 27, no1, p. 011007, 2018.