International Journal of Multidisciplinary Studies and Innovative Technologies

Volume : 3
Number: 1
Year: 2019
Pages: 1-4

Skin Segmentation by Using Complex Valued Neural Network with HSV Color Spaces*

Eda Çapa Kızıltaş[1*], Ayşenur Uzun[2] and Ersen Yılmaz[2]

[1]*Borçelik Çelik Sanayi Ticaret A.Ş., Bursa, Turkey*
[2]*Electrical-Electronic Engineering Department, Uludag University, Bursa, Turkey*
[*]*Corresponding author: eckiziltas@borcelik.com*

Abstract – Nowadays, digital image processing is useful tool in daily life for security, surveillance and artificial intelligence applications. Mainly in nudity alerts and face detection, skin segmentation is widely preferred method due to its simple background. The main problem in skin segmentation is skin color variation that result of different percentage of pigments, number and size of melanin particles in human-being. In literature, there are rule based and hybrid models for skin segmentation, however rule-based algorithm are not enough to overcome skin variety. As hybrid mode Complex valued neural network (CVNN) and color space transformation is applied to Skin Segmentation Database from UCI Learning Repository that is collection of different age and race group human's skin samples in RGB format.

Keywords – Skin Segmentation, Complex valued neural network (CVNN), Machine Learning

I. INTRODUCTION

Skin segmentation is a helpful method in image filtering to select skin and non-skin regions through an image [1]. Skin segmentation is widely used in image and video processing [2,3] for detection and tracking applications [4], recognition of gestures [5]. The main reason of being useful image filter is that skin segmentation is affordable computational time and independency of posing [6]. Although human skin color is differentiable than other objects, the variation in color scale among people and different illumination due to external factors complicates the process.

Ethnicity takes role in skin color variation in human by arising different gene pool and environmental conditions. Amount and distribution of heritable pigments as melanin results in alteration of darker skin to lighter skin [7]. Variation in color effects segmentation process negatively. Moreover, illumination due to face orientation, camera specification and inadequate lightening is degrading effect in segmentation classifiers [1].

The skin segmentation dataset is retrieved from UCI Machine Learning Repository. (http://archive.ics.uci.edu/ml/datasets/Skin+Segmentation). Dataset is combination of skin textures from face images that is collected both from Color FERET Image Database (http://face.nist.gov/colorferet/request.html) and Productive Aging Laboratory Database (http://agingmind.utdallas.edu/facedb/). Face images belong to people from different age, gender and race are strengthen dataset by providing variation. Dataset include 3 features of color information in RGB format and two classes as skin and non-skin - numbered as 1 and 2. Total sample number is 245057 that 50859 of them is skin textures.

In 2005 [6] studied to classify skin textures with machine learning approach by using ECU dataset. They applied Bayesian classifier and Multiplayer Perceptron with accuracies 89.79%, 89.49%. [8] classified face image dataset in SFA as skin and non-skin pixel. The proposed method for pattern classification was artificial neural network and 92.71% accuracy is achieved. In same study, they compared artificial neural network efficiency in the UCI skin segmentation dataset. Segmentation accuracy was 88.74%. After [8], [9] improved neural network model and raised accuracy to 99%. In 2018 neural network model was applied to dataset that is fed with hue, saturation, value (HSV) transformation rather than RGB, but accuracy was lower than earlier studies; 94.31% [10]. UCI dataset was classified with 3 different machine learning approach in [11]. Deep neural network was the most fitting method according to study with 97.32% accuracy in comparison to navie Bayes and decision tree which had accuracies 93.23% and 96.35%. [12] extracted skin color features of UCI database with RGB, HSV, YCbCr and CIELab color spaces and resulted in single feature vector to train Gaussian mixture model. The accuracy value for UCI dataset was 97.88%.

CVNN (complex valued neural network) is a special neural network application with complex value algebra. Especially for fields in complex nature like telecommunication, signal processing, speech processing, CVNN enforce the applicability of artificial neural network [13]. In addition, complex values decrease operation time, network size and speed training process with its two-dimensional geometry [14]. With respect to real valued neural network (RVNN), CVNN has better generalization character due to orthogonality of decision boundaries [13].

In our study Single Layered Complex Valued Neural Network is used to classify skin and non-skin textures. 10-fold cross validation is applied in order to validate method in dataset.

II. MATERIALS AND METHOD

A. *Complex Valued Neural Network*

Complex value algebra is applied to conventional neural network approach for CVNN. The basic perceptron illustration is shown in Figure 1.

**This study is extended version of the abstract paper "Skin Segmentation by Using Complex Valued Neural Network" that is presented in ISAS 2018-Winter, 2nd International Symposium on Innovative Approaches in Scientific Studies.*

Real valued inputs should be mapped to complex values. Phase encoding method is efficient way to represent a real value in complex geometry [15]. Phase encoding transforms real valued samples to the upper half of unit circle which is representation of all complex value with absolute value - length- 1. As a result, the phase transformation used in this study reveal unity complex numbers with angle between $[0 - \pi]$.

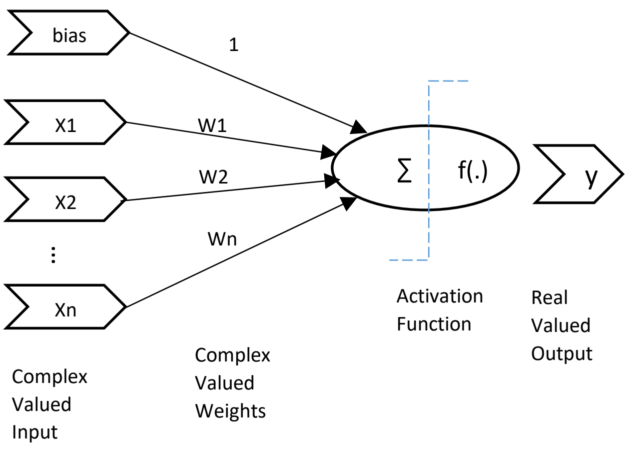

Fig. 1 Complex Valued Neuron [16-10]

For complex perceptron that has X input with n feature $[x_1 \quad x_2 \quad \cdots \quad x_n] \in R_n$ and Y output with m class $[y_1 \quad y_2 \quad \cdots \quad y_n] \in R_m$; all features are transformed with their extreme values; minimum and maximum points $-x_j \in [x_{tmin} \quad x_{tmax}]$ for $x_{tmin}, x_{tmax} \in$ R.

$$\varphi = \frac{\pi(x-a)}{(b-a)} \tag{1}$$

Complex transformed of x_j *is* c_j where;

$$c_j = e^{i\varphi} = \cos\varphi + i\sin\varphi \tag{2}$$

As a result, complex inputs of perceptron become $[c_1 \quad c_2 \quad \cdots \quad c_n]$.

After encoding, complexed inputs are weighted with complex weight parameter of perceptron; w_{kj} from j^{th} input to k^{th} neuron. The overall weighted inputs and complex valued bias of neuron; θ_k; are summed. z_k is the complex output of perceptron before activation function.

$$z_k = \sum_{j=1}^{m} w_{kj}\, c_j + \theta_k \tag{3}$$

Complex algebra behind the Equation 3 can be summarized as;

$$z_k = \left(\sum_{j=1}^{m} w_{kj}^R c_j^R - w_{kj}^I c_j^I + \theta_k^R\right) +$$

$$i\left(\sum_{j=1}^{m} w_{kj}^I c_j^R - w_{kj}^R c_j^I + \theta_k^I\right) \tag{4}$$

for $\theta_k = \theta_k^R + i\,\theta_k^I$; $w_{kj} = w_{kj}^R + i\,w_{kj}^I$ and $c_k = c_k^R + i\,c_k^I$

In CVNN approach, the activation function of perceptron is a bit frustrating because of gradient-descent back propagation algorithm which yields to update weights of perceptron in each iteration for fitting network. Activation functions are used to bound output, however; according to Liouville's theorem a complex-valued function cannot be both differentiable and bounded [15]. In other words, conventional activation functions are not suitable to be used with complex values for back propagation learning method. In literature two activation functions are proposed that map complex valued outputs to real valued classes [15].

For summed weighted inputs of perceptron,

$$z_k = z_k^R + i\,z_k^I \tag{5}$$

two proposed activation functions are as;

$$f_{C\to R}(z_k) = \sqrt{(f_R(u))^2 + (f_R(v))^2} = y_k \tag{6}$$

$$f_{C\to R}(z_k) = (f_R(u) - f_R(v))^2 = y_k \tag{7}$$

where, $f_R(x)$ is the sigmoid function;

$$f_R(x) = 1/(1+\exp(-x)) \tag{8}$$

In gradient-descent learning rule step, the weights and bias of neuron are updated for minimizing error function;

$$E = \left(\frac{1}{2}\right)\sum_{k=1}^{n}(o_k - y_k)^2 = \left(\frac{1}{2}\right)\sum_{k=1}^{n}(e_k)^2 \tag{9}$$

where, o_k is the expected output of neuron.

Minimizing process is act of differentiating function. In back propagation learning method, partial derivative of error function is stepped by learning rate to search minimum point of curve with trend changes at the specified point. Different from real valued network, partial derivatives of error function are calculated for real and imaginary part respectively.

For biasing of perceptron, partial derivative – delta of bias – is as;

$$\theta_k = \theta_k + \Delta\theta_k \tag{10}$$

Similarly, for weight values;

$$w_{kj} = w_{kj} + \Delta w_{kj} \tag{11}$$

which yields;

$$\Delta w_{kj} = \overline{c_j}\Delta\theta_k \tag{12}$$

where, $\overline{c_j}$ is the complex conjugate of j^{th} complex input.

For activation function in Equation 6; the derivative of error function is;

$$\Delta\theta_k^R == \eta e_k \frac{f_R(z_k^R)}{y_k} f_R'(z_k^R) \tag{13}$$

$$\Delta\theta_k^I == \eta e_k \frac{f_R(z_k^I)}{y_k} f_R'(z_k^I) \tag{14}$$

For equation 7;

$$\Delta\theta_k^R = 2\eta e_k (f_R(z_k^R) - f_R(z_k^I)) f_R'(z_k^R) \tag{15}$$

$$\Delta\theta_k^I == 2\eta e_k (f_R(z_k^I) - f_R(z_k^R)) f_R'(z_k^I) \tag{16}$$

In both function $f_R(x) = 1/(1+\exp(-x))$ is sigmoid activation function so the derivative of $f_R(x)$with respect to x;

$$f_R'(x) = f_R(x)\big(1 - f_R(x)\big) \tag{17}$$

B. Colour Space Transformation

Human eye tends to sample reflected light – colour of images – in three bands that fit to red, green, blue light wavelengths. However, in digital imaging process various colour space transformation is used to enhance convenient property of light in distinguishing features. Hue, intensity, brightness, chroma are some of these properties.

To eliminate nonuniform illumination in images, normalization over merits of colour (Nrgb) is done [17].

$$\Delta_R(c) = \frac{I_R(c)}{I_R(c) + I_G(c) + I_B(c)} \tag{18}$$

Where $[I_R(c), I_G(c)\ , I_B(c)]$ is the c^{th} pixel RGB values and $[\Delta_R(c), \Delta_G(c)\ , \Delta_B(c)]$ is the Nrgb form of c^{th} pixel.

Hue, Saturation and Value (HSV) colour space is non-linear transformation of RGB colour space that is closed to human perception [18] [19]. Hue can be assumed as colour portion that is represented as angle. Saturation is depth of colour related with gray amount. Value is brightness of pixel [18].

$$H(c) = \cos^{-1} \frac{0.5*\big(2I_R(c) - I_G(c) - I_B(c)\big)}{\sqrt{(I_R(c) - I_G(c))^2 - \big(I_R(c) - I_B(c)\big)\big(I_G(c) - I_B(c)\big)}} \tag{19}$$

$$S(c) = \frac{\max\big(I_R(c), I_G(c), I_B(c)\big) - \min\big(I_R(c), I_G(c), I_B(c)\big)}{\max\big(I_R(c), I_G(c), I_B(c)\big)} \tag{20}$$

$$V(c) = \big(I_R(c), I_G(c), I_B(c)\big) \tag{20}$$

where, $[I_R(c), I_G(c)\ , I_B(c)]$ is the c^{th} pixel RGB values and $[H(c), S(c)\ , V(c)]$ is the HSV form of c^{th} pixel [20].

C. k-fold Cross Validation (CV)

In machine learning applications, dataset is separated into two parts; one for training model and other for testing performance against the overfitting. The main challenge in separation is ensuring randomness while representing all occurrence in both train and test data. CV is a strong statistical tool which let all samples to both train and test model. Samples are divided into k groups in $k - fold$ cross validation and each iteration one of group is held as test subset while remaining groups are used to train network [16]. The overall performance of model over dataset is average of k iterations.

III. Results

In this study; image pixels of both skin and non-skin objects in RGB image format is classified with single layered CVNN.

The CVNN model has 3 complex inputs by phase encoding and 2 real valued output for classification vector that $[0 \quad 1]$ is skin class and $[1 \quad 0]$ is non- skin class.

The dataset is divided into 10 groups randomly for 10-fold cross validation and each group is used for testing data in each turn. Input features are normalized to before phase encoding.

CVNN model is trained with back-propagation updating by using gradient descent as learning algorithm. The learning rate used in this study is 0.0001. The maximum iteration of model is 1000 cycle.

At the beginning of training steps, all real and complex parts of bias and weights of neurons are initialized randomly between 0 – 0.1.

The activation function used in complex perceptron that mapping neuron outputs to real valued class vectors is;

$$f_{C\to R}(z_k) = (f_R(u) - f_R(v))^2 = y_k \tag{18}$$

In this study skin segmentation dataset is classified with average root mean square error 0.169 and 98.60% average accuracy in 10-fold cross validation with HSV colour transformation.

Table 1. 10-fold RMSE and MAPE for HSV

	RMSE	MAPE
1st	0.169	0.986
2nd	0.165	0.987
3rd	0.168	0.986
4rd	0.170	0.986
5th	0.173	0.985
6th	0.166	0.986
7th	0.165	0.987
8th	0.169	0.986
9th	0.173	0.985
10th	0.169	0.986

Table 2. The confusion matrix for 3rd fold.

Class/ Observed	3rd fold	
	1	2
1	5446	115
2	222	18723

With help of Nrgb transformation, skin segmentation dataset is classified with average root mean square error 0.704 and %77.56 average accuracy in 10-fold cross validation.

Table 3. 10-fold RMSE and MAPE for Nrgb

	RMSE	MAPE
1st	0.703	0.779
2nd	0.711	0.7555
3rd	0.699	0.791
4rd	0.678	0.81
5th	0.715	0.7512
6th	0.706	0.775
7th	0.705	0.7712
8th	0.701	0.7923
9th	0.709	0.761
10th	0.708	0.7698

Table 4. The confusion matrix for 3rd fold.

Class/ Observed	3rd fold	
	1	2
1	5446	115
2	222	18723

IV. CONCLUSION

The success of the model for classifying the dataset is higher in accuracy for HSV colour transformation rather than Nrgb. For further study, this model can be applied to skin detection or face detection process. Moreover, other skin segmentation of face images dataset as SFA can be used to validate success of constructed single layered CVNN.

ACKNOWLEDGMENT

This study is extended version of the abstract paper “Skin Segmentation by Using Complex Valued Neural Network” that is presented in ISAS 2018-Winter, 2nd International Symposium on Innovative Approaches in Scientific Studies.

REFERENCES

[1] Mahmoodi, Mohammad Reza. "High Performance Novel Skin Segmentation Algorithm for Images With Complex Background." *arXiv preprint arXiv:1701.05588 (2017).*

[2] J R.-L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, “Face detection in color images”, *IEEE Trans. PAMI, vol. 24, no. 5, pp. 696-706, 2002..*

[3] R. Schumeyer and K. Barner, “A color-based classifier for region identification in video”, *in Proc. of Visual Communications and Image Processing, vol. 3309, 1998, pp. 189-200..*

[4] Bhatt, Rajen B., et al. "Efficient skin region segmentation using low complexity fuzzy decision tree model." *India Conference (INDICON), 2009 Annual IEEE. IEEE, 2009.*

[5] A. Gupta, V. K. Sehrawat, and M. Khosla. "FPGA based Real T im e Human Hand Gesture Recognition System,." *Procedia Technology 6 (2012): 98-107.*

[6] Phung, Son Lam, Abdesselam Bouzerdoum, and Douglas Chai. "Skin segmentation using color pixel classification: analysis and comparison." *IEEE transactions on pattern analysis and machine intelligence 27.1 (2005): 148-154.*

[7] Costin, Gertrude-E., and Vincent J. Hearing. "Human skin pigmentation: melanocytes modulate skin color in response to stress." *The FASEB journal 21.4 (2007): 976-994.*

[8] Casati, Joao Paulo Brognoni, Diego Rafael Moraes, and Evandro Luis Linhari Rodrigues. "SFA: A human skin image database based on FERET and AR facial images." *IX workshop de Visao Computational, Rio de Janeiro*. 2013.

[9] Yasar, Ali, and Ismail Saritas. "Determination of whether Skin or non Skin from the Color Pixels Using Neural Network." *Managing Intellectual Capital and Innovation for Sustainable and Inclusive Society: Managing Intellectual Capital and Innovation; Proceedings of the MakeLearn and TIIM Joint International Conference 2015*. ToKnowPress, 2015.

[10] Lal, Anisha M., D. Aju, and Mradima Sharma. "SKIN SEGMENTATION OF RGB IMAGES AND ADAPTIVE RECOGNITION USING HSV COLOR MODEL."

[11] Dastane, Tejas, et al. "An Effective Pixel-Wise Approach for Skin Colour Segmentation-Using Pixel Neighbourhood Technique." *International Journal on Recent and Innovation Trends in Computing and Communication* 6.3 (2018): 182-186.

[12] Jaisakthi, S. M., and S. Mohanavalli. "Skin Segmentation using Ensemble Technique." Research Journal of Applied Sciences, Engineering and Technology 9.11 (2015): 963-968.

[13] Nitta, Tohru. "Orthogonality of decision boundaries in complex-valued neural networks." *Neural Computation 16.1 (2004): 73-97.*

[14] Peker, Musa, Baha Sen, and Dursun Delen. "A novel method for automated diagnosis of epilepsy using complex-valued classifiers." *IEEE journal of biomedical and health informatics 20.1 (2016): 108-118.*

[15] Amin, Md Faijul, and Kazuyuki Murase. "Single-layered complex-valued neural network for real-valued classification problems." *Neurocomputing 72.4-6 (2009): 945-955.*

[16] Yılmaz, Ersen. "Fetal State Assessment from Cardiotocogram Data Using Artificial Neural Networks." *Journal of Medical and Biological Engineering 36.6 (2016): 820-832.*

[17] Chude-Olisah, Chollette C., et al. "Illumination normalization for edge-based face recognition using the fusion of RGB normalization and gamma correction." Signal and Image Processing Applications (ICSIPA), 2013 IEEE International Conference on. IEEE, 2013.

[18] Sural, Shamik, Gang Qian, and Sakti Pramanik. "Segmentation and histogram generation using the HSV color space for image retrieval." Image Processing. 2002. Proceedings. 2002 International Conference on. Vol. 2. IEEE, 2002.

[19] Yu, Hui, et al. "Color texture moments for content-based image retrieval." Image Processing. 2002. Proceedings. 2002 International Conference on. Vol. 3. IEEE, 2002.

[20] Shaik, Khamar Basha, et al. "Comparative study of skin color detection and segmentation in HSV and YCbCr color space." Procedia Computer Science 57 (2015): 41-48.