Merging Segmentation Capabilities of Independent Classifiers

Sidra Gul

Abasyn University
Peshawar, Pakistan

Laiq Hassan

University Of
Engineering and
Technology, Peshawar,
Pakistan

Kashif Ahmad

University Of
Engineering and
Technology, Peshawar,
Pakistan

Kamal Haider

Gandhara Institute of
Science and Technology,
Peshawar, Pakistan

ABSTRACT

In the modern era of computer and technology, images and videos play a vital role. Therefore, there is always a need for robust skin detection system in order to cope with the intolerable and objectionable contents. In this paper, an efficient method has been implemented for skin detection, which detects the skin in different images under different environmental conditions. We have used the two machine learning approaches i.e. Random Forests and Multilayer perceptron for skin detection. We have also then combined the results of these two approaches used. We have used total of 554 images in our experiments.

General Terms

Skin detection, recognition, tracking, images, and classification

Keywords

Random Forests, Multilayer Perceptron, color spaces, bitwise operators, merging.

1. INTRODUCTION

Images and videos are an important part of our everyday routine, whether you need a study tutorial or entertainment. When there is an immense interaction of human and computer for different types of videos, either intentionally or intentionally one can come up with the contents that are prohibited. Therefore, there is always a need for implementing techniques that could cope with such situations. Skin detection is one of the many techniques which are used to block prohibited contents and give the user only with the data which he needs. Skin detection may seem easy but it is a challenging task.

Skin detection is the process of finding skin-color pixels and regions in an image or video. In images and videos, skin color is an indication of the existence of humans in media. In one of the many applications, detecting skin-color regions was used to identify nude pictures on the internet for filtering the contents [1]. Skin detection was also used in the applications such as to detect anchors in TV news videos for the sake of video automatic annotation, archival and retrieval [2, 3]. A skin detector typically transforms a given pixel into an appropriate color space and then uses a skin-color classifier to label the pixel whether it is skin or non-skin. There has also work been done on the fusion of two different methods for skin detection [4].

Skin detection is also used for other purposes like face tracking, gesture recognition, security reasons and content based image retrieval systems [2, 3, 5]. In all such application, the desired object is detected by deleting the background elements. This is done by detecting only the skin regions in the image, such as hand, face etc. Different techniques exist for detection of skin. One of the most commonly used techniques for detecting skin region is the color based skin detection [6]. In color-based skin detection, a fixed boundary is defined for each color component of a particular color space to separate the skin region from the non-skin region. The different color spaces that exit are RGB, HSV [7] and YCbCr [8] etc. Each of these color spaces has their own characteristics and calculations. In addition to the color spaces, color constancy is another factor in skin color detection. There is difference between, how humans and computers perceive colors.

Machine learning algorithms have been widely used in the field of image processing [9]. These approaches include multilayer perceptron, Bayesian network, random forests, fuzzy based skin detection [10, 11] and self-organizing map [12]. All of these algorithms have the mechanism to learn from the input data which we called as training. Different machine learning algorithms have their own learning procedures which make them different from each other on the basis of efficiency and performance.

In our approach we have considered the RGB color space along with machine learning algorithms for skin color detection. The two machine learning algorithms we have used are random forests [13] and multilayer perceptron or neural networks [14]. We have used random forests and multilayer perceptron due to its efficiency and performance as compared to other approaches. Our main objective, after achieving the results of random forest and multilayer perceptron, is to combine these two results and then analyze the final results. There are two possibilities of combining the results. This can increase the system's efficiency, and the result can be better. On the other hand there is also a possibility that after merging the results are not better than the separate results of the two approaches used.

The remaining paper is divided into five sections. Section 2 is about Random forests details. Section 3 explains Multilayer Perceptron. Section 4 is about the merging and fusion of the two different approaches. Section 5 contains the experiments and results; section 6 is about the conclusions of the evaluated results.

2. RANDOM FOREST

Now day's random forests are widely used due to its huge popularity in the field of computer applications because it has easily training algorithms. These were introduced by Tin Ho

[15]. The type of learning used by random forest is called supervised learning [16]. Random forests have been widely used for the detection of human skin in an image [17].

Random forest is a collection of large number of trees. These trees depend on the values of a random vector [18]. This vector is given to each tree in the random forest. Each tree evaluates the input vector and gives it result independently. The results of each tree are called as vote. The word "vote" has been used because more than one tree gives same results for the input vector. Then the number of votes is counted at the end. The result, for which more trees vote, is chosen at the end. Let there are N cases in the training set. Then Random forest is trained for all the N cases in the training set in a random fashion, one by one and by replacing the original data. If there are large number of variables, run all the data once on the random forests and then choose only the important data to be run on the random forest for the second time.

The construction of the tree is such that if there are total of M variables in the classifier, then m input variables are chosen to find the decision at the node, whereas $m \leq M$. The variable m is chosen on random bases. The next split is calculated, which is based on the variable m.

3. MULTILAYER PERCEPTRON

Multilayer perceptron consists of two or more than two layers between an input and output layer, such that they form a network like human sensory neurons. They are also called as feed forward artificial neural networks. Neural networks have played its role in many fields of computer science that includes underwater sonar current, speech recognition and many more. It has also been used in the prediction of globular proteins. Another very important application of neural networks is its use in statistical models such as, its use in the financial problems, sales and marketing etc.

Artificial neural networks has also shown great efficiency in the field of image processing and pattern recognition, such as to recognize face [19] or any character[20]. One can create an artificial neural network in a computer by simply simulating the neural model of the human brain. Then we can make this network learn through training process also called supervised learning [16], which can then solve different kinds of problems. The training process depends upon the kind of problem and the data sets provided to the learning process.

There are different kinds of artificial neural networks, i.e. multilayer perceptron, radial basis function networks and kohonen neural networks. The approach which we have implied is the multilayer perceptron. In perceptron neural networks are formed in a number of hierarchies of layer, Fig1.

The very first layer is the input layer, through which we give input to the network, whereas the last layer is the output layer. Between these two layers are many hidden layers. The number of these hidden layers depends on the type of application and desired performance to be achieved.
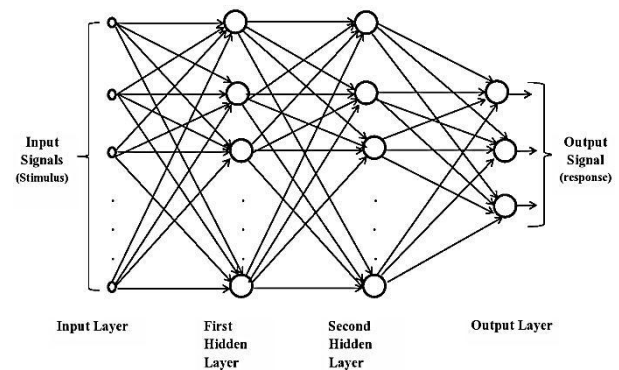

Fig 1: Multilayer Perceptron Neural Network with two hidden layers

A simple neural network can be given by equation 1,

$$out_k = \sum_{i=1}^{m} w_{ki}\ x_i \qquad (1)$$

In the above equation out_k is the desired output, w_{ki} are the weights and x_i is the input. m is the number of hidden layers. This equation represents one layer in a hierarchy of many layer of neural networks, in which each weight of the perceptron is multiplied with its input and then the sum of all these is provided as an input to the next adjacent layer in the hierarchy. Similarly the accumulated values are passed along the series of weighted connections and at the end to the network's output layer. The network is given the initial weights at first, and the data is then fed into the network.

4. FUSION OF THE TWO APPROACHES

Fusion is strategies that combines or integrate the results of two different methods into one single representation. In our case of skin detection, fusion is done in order to increase the effectiveness and robustness of the two skin detection algorithms used. Both the models will produce their own results.

The purpose of fusion can be achieved by using different methods. One of the fusion strategies is to take the product of the two algorithms [4]. Other approaches include bitwise operators i.e. AND, OR operators [21].

Graph cut is another method, use to combine two different approaches [22].Graph cuts have proven to be one of the very useful tools in vision and graphics [23].

5. EXPERIMENTAL RESULTS

In order to train the machine learning algorithms, we have performed different experiments by providing the data sets to them. Different results for the machine learning algorithms are accumulated which are discussed in this section.

5.1 Data Sets

The data sets of our experiments are obtained from the skin detection experiments in [24]. This data set contains two types of images, one is the original set of images Figure 2 and the other one is mask images Figure 3. The set of mask images have been used to train the machine learning

algorithms first, such that it learns to differentiate skin pixels from non-skin pixels. As discussed earlier, this type of learning is called supervised learning. We can then provide the original set of images to machine learning algorithms, once they are trained.

Fig 2: Original images from data sets

Fig 3: Mask images from data sets

5.2 Algorithm

In our algorithm we have used the RGB color space to detect skin region in colored images. The RGB color space equations are given by:

$$n_r = \frac{R}{R+G+B} \tag{2}$$

$$n_g = \frac{G}{R+G+B} \tag{3}$$

$$n_b = \frac{B}{R+G+B} \tag{4}$$

The results of RGB are then passed to the two machines learning algorithms i.e. Random Forest section 2 and Multilayer Perceptron section 3, to detect skin in different images. The machine learning algorithms are trained first with the training dataset. After training we provided them the original set of images.

After obtaining the results of random forests and multi- layer perceptron, the next step is to combine the results of the two algorithms. The operators being used in the experiments are bitwise operators i.e. bitwise multiplication, addition, AND, OR [21] and subtraction. In order to define different bitwise operations on the images, we have converted these images to binary first.

5.3 Results and Discussions

In Figure 4, images from left to right, Figure 4 (a) shows images from the original set of images, Figure 4(b) shows the results of random forest and Figure 4(c) shows the results of MLP for these images. If we compare the results of

the two algorithms i.e. MLP and Random forest, we can see that the results of MLP are much better than the Random forest results, due its highly trained procedures.
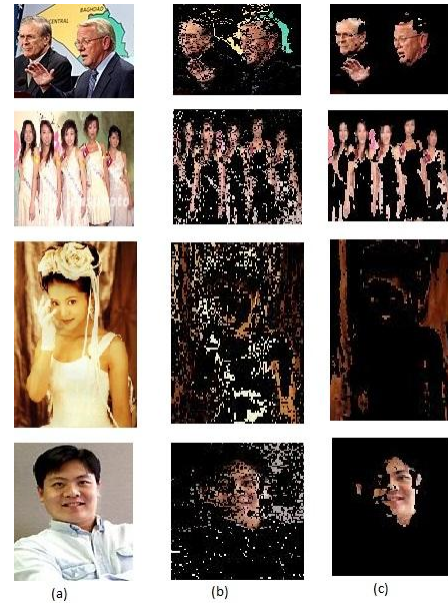

Fig 4: (a) Original images (b) Random Forests results (c) Multilayer perceptron results

In order to perform bitwise operation, the images in figure 4 are converted in to binary images, as shown in Figure 5 (a) and Figure 5 (b).
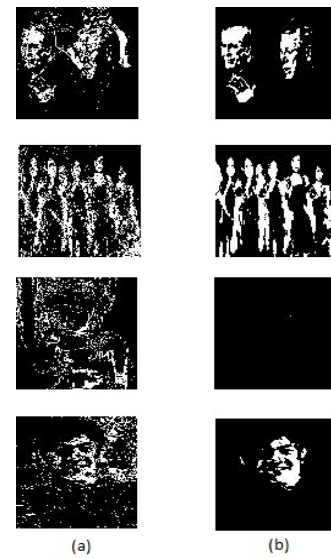

Fig 5: (a) Random Forests (b) Multilayer perceptron

Figure 6(a) shows the results of addition of the two algorithms used and Figure 6(b) shows the bitwise multiplication results. If we see the addition results, it shows that some additional information have been added in the image, due to which the results are not clear. In multiplication results, however only those areas are visible, which are detected as skin region in both Random Forest and MLP.

Fig 6: (a) Addition of Random Forest and MLP (b) Multiplication of Random Forest and MLP

Figure 7 (a) and Figure 7(b) shows the bitwise operations of AND and OR operators. As we can see none of these results are satisfactory. However, the results of bitwise ORing are somewhat better than the bitwise ANDing.
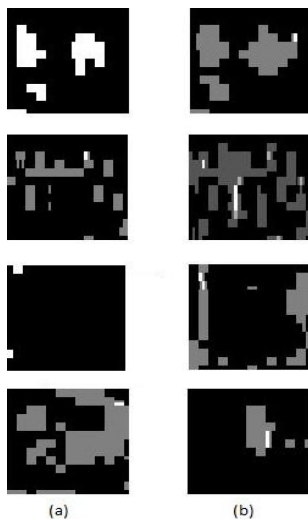

Fig 7: (a) Bitwise Anding of Random Forest and MLP (b) Bitwise Oring of Random Forest and MLP

Figure 8(a) shows the result of subtraction of Random Forest from MLP, whereas 8(b) shows the results of subtraction of MLP from Random Forest. The results from these operations give very little information about the skin region detected by both MLP and Random forest.
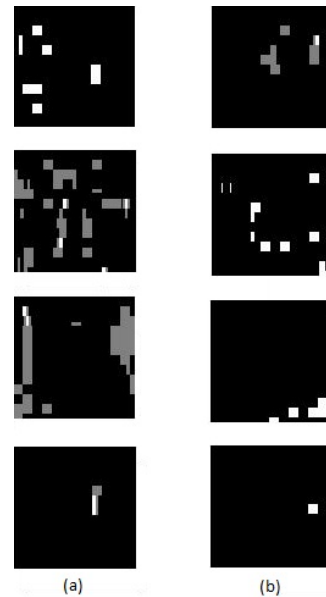

Fig 8: (a) Subtraction of Random Forest from MLP (b) Subtraction of MLP from Random Forest

6. CONCLUSIONS

In this paper, we concluded that the results of merging of the machine learning approaches i.e. random forest and multilayer perceptron, using bitwise operators are not that much satisfactory as are the individual results of the two approaches used.

If we see the individual results given by MLP and Random Forest, we can find that they have clear results than the merged results of Random Forest and MLP.

In some cases i.e. bitwise multiplication and bitwise ORing, the results are somewhat closed to the results of the individual machine learning approaches. The result of bitwise multiplication is also similar to that of the bitwise Anding. However the overall results do not give much information about the skin region in the image as do the individual results.

The individual results, however also includes some of the background information, but are better than the merged results. Therefore the bitwise operators cannot be the best choice for merging the segmentation results in case of machine learning approaches.

The other possible solution for the purpose of merging can be the graph cut method, which can be the future work in this case.

7. ACKNOWLEDGMENTS

REFERENCES

[1] S. Mofaddel, M.A.; Sadek, "Adult image content filtering: A statistical method based on multi-color skin modeling," Computer Technology and Development (ICCTD), 2010 2nd International Conference on, pp. 682,686, 2010.

[2] C. S. Chan, H. Liu, and D. J. Brown, "Recognition of human motion from qualitative normalised templates," J. Intell. Robot. Syst., vol. 48, no. 1, pp. 79–95, 2007.

[3] N. Kubota and K. Nishida, "Perceptual control based on prediction for natural communication of partner robot,"

IEEE Trans. Ind. Electron., vol. 54, pp. 866–877, Apr. 2007.

[4] W. R., Tan, C. S., Chan, P. Yogarajah, and J. Condell, "A fusion approach for efficient human skin detection," IEEE Transactions on Industrial Informatics, vol. 8, no. 1, pp. 138–146, 2012.

[5] P. Vaddakkepat, P. Lim, L. De Silva, L. Jing, and L. L. Ling, "Multi- modal approach to human-face detection and tracking," IEEE Trans. Ind. Electron., vol. 55, pp. 1385–1393, Mar. 2008.

[6] H. C. Vijay, Lakshmi and S. Patikulakarni, "Segmentation algorithm for multiple face detection in color images with skin tone regions using color spaces and edge detection techniques," International Journal of Computer Theory and Engineering., vol. 2, no. 4, pp. 1793–8201, 2010.

[7] V. A. Oliveria and A. Conci, "Skin detection using hsv color space".

[8] V. Powar, A. Jahagirdar, and S. Sirsika, "Skin detection in ycbcr colorspace," International Journal of Computer Applications, pp. 00975–8887, 2011.

[9] M. Emre, Celebi, H. A., Kingravi, Y. Alp, Aslandogan, and W. V., Stoecker, "Detection of blue-white veil areas in dermoscopy images using machine learning techniques," Medical imaging 2006:Image Processing, vol. 6144, 2006.

[10] A. Olano, A. Soria-Frish, and R. Verschae, "Robust skin detection based on the fuzzy integral," 2005.

[11] Y. Chahir and A. Elmoataz, "Skin-color detection using fuzzy cluster- ing," 2005.

[12] D. Brown, I. Craw, and J. Lewthwaite, "A som based approach to skin detection with application in real time systems".

[13] F. Livingston, "Implementation of breiman's random forest machine learning algorithm," ECE591Q Machine Learning Journal Paper., 2005.

[14] G. Peter, Zhang, "Neural networks for classification:a survey," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICSPART C: APPLICATIONS AND REVIEWS, vol. 30, no. 4, pp. 451–462, 2000.

[15] T. K. Ho, "Random descision forests," in ICDAR, pp. 278–282, 1995.

[16] V. Gullapalli, "A comparison of supervised and reinforcement learning methods on a reinforcement learning task," Intelligent Control, 1991., Proceedings of the 1991 IEEE International Symposium on, pp. 394,399,1991.

[17] R. Khan, A. Hanbury, and J. Stoettinger, "Skin detection: A random forest approach," in Proc. Int. Conf. Image Process., pp. 4613–4616, 2010.

[18] L. Breiman, "Random forests," Mach. Learn, vol. 45, no. 1, pp. 5–32, 2001.

[19] O. S. M. El-Khamy, S.E.; Abdel-Alim, "Neural network face recognition using statistical feature and skin texture parameters," Radio Science Conference, 2001. NRSC 2001. Proceedings of the Eighteenth National, vol. 1, pp. 233–240, 2001.

[20] V. Kertesz, A.; Kertesz, "Dynamically connected neural network for character recognition," Neural Networks, 1992. IJCNN. International Joint Conference, vol. 3, pp. 672–676, Jun 1992.

[21] C. Amy, Doukim, J. Ahmad, Dargham, A. Chekima, and S. Omatu, "Combining nerural networks for skin detection," Signal and Image Processing: An International Journal, vol. 1, no. 2, 2010.

[22] P. Brain, L., M. Bryan, S., and C. Scott, "Geodesic graph cut for inter- active image segmentation," Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference, pp. 13161–13168, 2010.

[23] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-d image segmentation," International Journal of Computer Vision, 2006.

[24] Q. Zhu, C.-T. Wu, and Y.-L. W. Kwang Ting Cheng, "An adaptive skin detection," ACM International Conference on Multimedia, 2004.