Neural Network used in Energy Function of Content-aware Image Resizing Method to Preserve Skin Color Region

Tamara Arruda Pereira , Savio Lopes Rabelo ,
Saulo Freitas Oliveira , Francisco Nivando Bezerra

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE),
Eixo Tecnológico de Computação, Campus Maracanaú – CE – Brasil

`{tamaraarrudap, saviorabelo.ti, saulo.freitas.oliveira}@gmail.com`

`nivando@ifce.edu.br`

Abstract. *Seam Carving is a content-aware image resizing method capable of modifying the width or height to pictures preserving important elements. To evaluate the importance of each pixel this algorithm uses a energy function. The original algorithm can cause frequent object deformation in specific cases such as images containing people, because the energy function is not able to detect people. With this in mind, this paper presents a reformulation on the energy function of Seam Carving specific for preserving people in images. This energy function uses a neural network followed by morphological operations. For the experiments we used two dataset widely used for this type of problem RetargetMe and NRID. For comparison was used quality analysis metrics of resized images. The results show that the proposed method achieves superior performance in terms of visual quality via qualitative indices compared to the original algorithm. It also has outstanding results when compared to other resizing operators.*

1. Introduction

Currently there are several types of digital devices and resizing images is widely used so that images can adapt to these devices. An efficient image resizing method must have the ability to preserve essential objects after resizing. Traditional techniques, such as uniform scaling, warping, among others, preserve the visual elements but causes audible distortion if the ratio is different between input and output. Another way to resize an image is to crop it based in some window. However, such approach can potentially discard relevant content if the output resolution is significantly smaller than the input image [Senturk & Akgun 2017].

There are also methods that use approaches that minimize distortion of objects after resizing, known as sensitive methods to content, an example is the Seam Carving (SC) [Avidan & Shamir 2007] a resizing method which is concerned with preserving the image content. To evaluate the relevant image content, SC uses a energy function. Strategies such as methods based on gradients, saliency maps, segmentation, wavelet transform decomposition are some methods known in the literature that can be employed as energy functions [Oliveira et al. 2015]. However, previously mentioned methods have the disadvantage of frequent deformations objects in the image because the algorithm removes the paths of pixels containing low energy until the desired size of the image is reached without considering the visual effect.

Seam Carving presents recurring content distortions when applied to resizing images with people, it happens because detect people is a challenging problem due the non-rigid nature of the human body produces innumerable possible poses, besides the variation of the size due to the position of the camera and partial occlusions and camouflage of the object caused by disorderly backgrounds [Nguyen et al. 2016]. Features extracted from images such as skin color and texture have the advantage of not varying in rotation, size and partial occlusions.

This paper presents the method that uses skin color as an input argument for a neural network to classify pixels as skin or non-skin, and uses this classification in the energy function of Seam Carving to preserve people after resizing. The proposed method was compared to the original SC using evaluation metrics of dissimilarity and similarity based on more salient features and details.

For a better understanding of the study presented in this article, this was organized in 7 sections: Section 1 presents the introductory aspects on the theme, Section 2 describes the Seam Carving method; Section 3 describes skin detection; Section 4 presents the method proposed in this paper: Resizing with preservation of people using Seam Carving; Section 5 describes the experiments and the results achieved by the proposed method; Section 6 describes a comparison of the proposed method with the method that inspired this article, Seam Carving; and finally, the conclusion is presented in Section 7.

2. Seam Carving

An effective resizing method modifies the size (height or width) of an image and should be able to preserve important content whenever possible. SC is an image resizing algorithm that uses an energy function to determine the relevant content that must be preserved [Dong et al. 2009]. This energy function must be able to assign more energy to the pixels that stand out in the image and less energy to the imperceptible pixels that blend into the environment. The original SC algorithm uses the magnitude of the gradient as energy function (Equation 1), defined as a filter based on derivatives that highlights details and detects borders.

For each pixel in the image $\mathbf{I}$, the magnitude of the gradient vector is calculated by summing the modules of the derivatives in the vertical and horizontal direction and is given by:

$$e(\mathbf{I}) = |\frac{\partial}{\partial x}\mathbf{I}| + |\frac{\partial}{\partial y}\mathbf{I}|. \tag{1}$$

The SC algorithm inserts or removes seams (a path of connected pixels that cross the image from top to bottom, or from left to right and which has only one pixel in each line or column of the image). To reduce the size of the image while preserving its structure, the method focus on removing low energy pixels rather than those with high energy [Avidan & Shamir 2007]. The seam that produces less energy is considered a optimal seam and can be found using dynamic programming in linear complexity. The first step is to traverse the image from the second row to the last row and compute the cumulative minimum energy for all possible connected seams for each entry (x, y):

$$\mathbf{M}(x,y) = e(x,y) + min(\mathbf{M}(x-1,y-1), \mathbf{M}(x-1,y), \mathbf{M}(x-1,y+1)), \tag{2}$$

where $\mathbf{M}(x, y)$ is the cumulative minimum energy and $e(x, y)$ is the magnitude of the gradient at the position (x, y).

The minimum value of the last line is indicated at the end of the connected vertical junction at the end of this process. Thus, in the second step we backtrack from this minimum entry on $\mathbf{M}$ to find the path of the optimal seam [Shamir & Avidan 2009]. The definition of $\mathbf{M}$ for horizontal seams is similar.

Reducing the width of an image by n pixels boils down to applying the SC operator n times. That is, in each time step we find the optimal seam in the image, remove the pixels associated with it and repeat the process for n times.

Fig. 1 shows the SC method applied to an image containing a person. In order to resize the image, 25% of seams from the total image size were removed. In Fig. 1(c), we can see that the image got deformed after the seam removal. Such a situation usually happens when the energy function is not suitable for the problem.
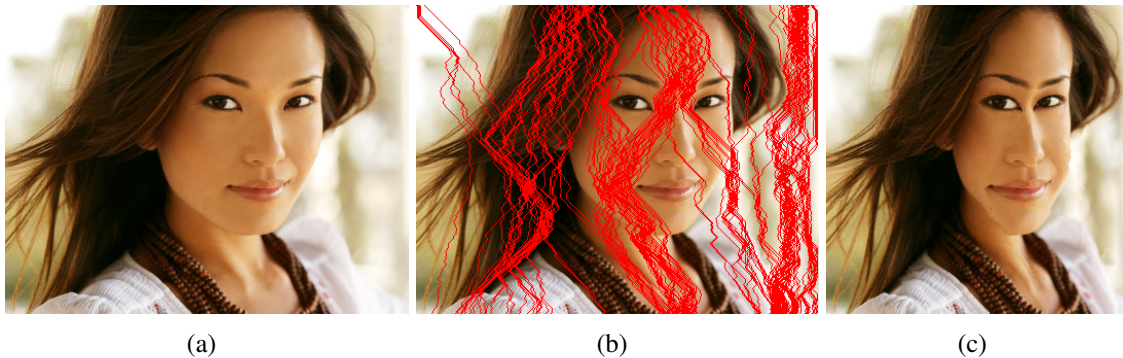

(a)   (b)   (c)

Figure 1. Use of the SC method. (a) Original image. (b) 25% of the vertical seams chosen to be removed. (c) Reduction of 25% of the width.

3. Skin Detection

Image segmentation is the process of partitioning an image into non-overlapping regions that are groups of homogeneous and connected pixels [Liu & Chung 2011]. Skin detection is an image segmentation technique that divides pixels into two groups: skin region (any exposed part of the human body that appears in the image, such as faces, shoulders, hands, legs) and non-skin region (everything that is not human skin is non-skin). This process takes into consideration some characteristics of the image, such as color or texture, in order to perform well.

Skin detection is used in a large range of image processing problems, such as face detection [Zhipeng et al. 2010], face recognition [Kang et al. 2016], gesture recognition [Han et al. 2009], content based image retrieval [Bhute & Meshram 2014], naked people detection [Perez et al. 2017], among others.

Skin color is an efficient property that is also robust and invariant to changes related to image rotation, scale and partial occlusion [Kakumanu et al. 2007]. However, detecting human skin by using color as a premise in complex images has proven to be a challenging problem because skin color can vary drastically due to illumination, race, aging, image conditions and complex background [Yuetao & Nana 2011].

A skin detector searches for pixels with colors that match the skin color model [Phung et al. 2003]. There are methods, such as artificial neural networks [Naji et al. 2018] that attempt to create skin detectors with the ability to learn from a set of training data without creating an explicit model of skin color. The output $\phi(x, y)$ of the neural network for pixel classification is given by:

$$\phi(x, y) = \begin{cases} 1, & \text{if } \mathbf{I}(x, y) \in \text{skin color,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

4. Preserving skin regions in Seam Carving

We propose a new content-aware image resizing method that focus on the preservation of people by using a new energy function for SC. The method is called Resizing with People Preservation using Seam Carving (RPP-SC) and combines the SC algorithm with a neural network capable of classifying pixel in two classes: skin and non-skin. To do so, the energy function is changed and the value $+\infty$ is assigned to all pixels classified as skin. The other pixels follow the procedure applied in [Avidan & Shamir 2007], which uses the magnitude of the gradient as energy function (Equation 1).

We replace the energy value for the pixels classified as skin color by $+\infty$. Thus, the new energy function e_ϕ for a pixel (x, y) is given by:

$$e_\phi(x, y) = \begin{cases} +\infty, & \text{if } \phi(x, y) = 1, \\ e(x, y), & \text{otherwise.} \end{cases} \tag{4}$$

The use of the energy functions presented in Equation 4 is shown in Fig. 2. Fig. 2(a) is the original image; Fig. 2(b) is the result after applying the energy function e of SC, the lighter the region, the more energy and more relevant; Fig. 2(c) is the classification of the pixel using the ϕ function; and Fig. 2(d) shows the image after using the energy function e_ϕ of RPP-SC, in which the pixels classified as skin are more evident compared to Fig. 2(b).
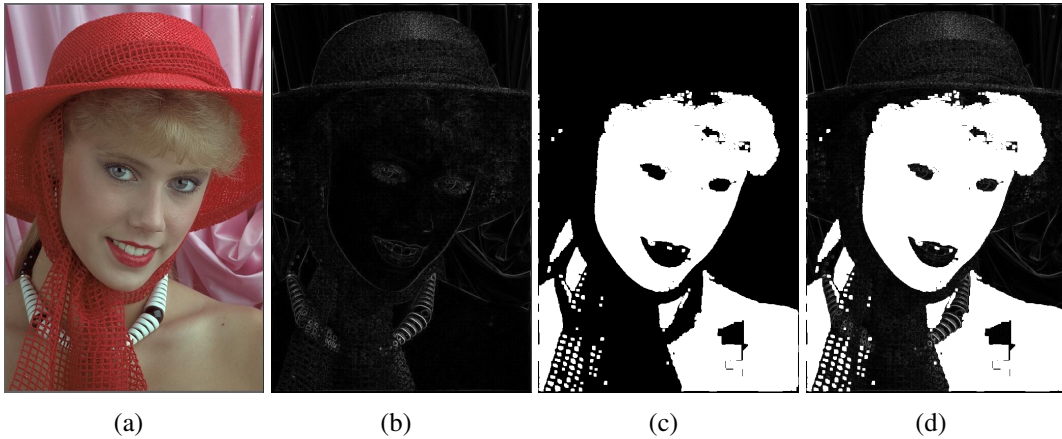

| (a) | (b) | (c) | (d) |

Figure 2. Energy functions. (a) Original image. (b) Energy function e. (c) Energy function ϕ. (d) Energy function e_ϕ.

The Fig. 3 shows the steps of the RPP-SC. First, we perform a pre-processing in the training database for the remove of ambiguity, then this data is passed as a parameter for network training. After the network has been trained and evaluated, the images that will be resize are passed to the network for skin detection, then a morphological operation is performed using a closure to remove small structures, with a square as a structuring element of size 5. Finally, the energy function e_{DP} is applied and the less important seams are removed, resulting in the resized images.
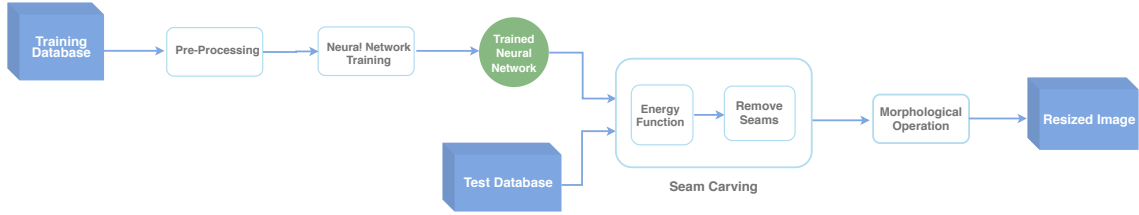

Figure 3. Stage for RPP-SC.

For the pixel classification as skin and non-skin, the chosen neural network was the Multilayer Perceptron (MLP) [Rumelhart et al. 1988]. The choice of database for the neural network training is important to ease problems such as race and aging. The SFA [Casati et al. 2013] database of human skin images was chosen because it is a data set that contains a diverse collection of images. The SFA was built based on facial images of FERET [Phillips et al. 1998] and AR [M. Martinez & Benavente 1998] databases, the SFA has 3354 skin samples and 5590 non-skin samples.

Color values for the RGB, HSV and YCbCr systems were used as attributes of the MLP input layer for skin classification, the same approach used in [Hajiarbabi & Agah 2015]. All data is preprocessed to remove ambiguities. Pixels that had the same values between skin and non-skin were assigned to the category of skin.

As skin color undergoes illumination variations, some approaches such as [Yuetao & Nana 2011] and [Baskan et al. 2002] assume that the chrominance components of skin tone color are independent of the luminance component. Therefore, the illumination channel is scorned and a two-dimensional color space is chosen instead of a three-dimensional color space. Thus, the HS can replace the HSV color space [Baskan et al. 2002] and the YCbCr [Habili et al. 2004] can be replaced by CbCr, discarding the luminance of these models. In [Moallem et al. 2011, Naji 2013], the authors assume that such approaches show poor performance when applied to complex images due to the loss of some color information when a colored pixel is expressed in two-dimensional rather than three-dimensional space. Ignoring any color information affects the accuracy of the system. For comparison purposes, this work applies both techniques, i.e. both the three-dimensional and two-dimensional color spaces are considered in our experiments.

5. Experiments

This section describes the experiments and the results achieved by the proposed method in this article. For all experiments, a computer with the following configuration was used: Intel (R) Core(TM) i7-6500U 2.5 GHz processor with 8 GB of RAM and was

executed the Windows 10. MATLAB R2017a (developed by MathWorks) was used for implementation.

5.1. Skin detection experiments

For skin detection experiments: about 60% of the samples were used for training, 20% for validation, and 20% for testing. MLP was trained with different numbers of neurons in the hidden layer, ranging from 6 to 40. For most networks, having 34 nodes in the hidden layer produced better results compared to another number of neurons. All experiments were repeated for 10 independent rounds.

Seven different type colorspaces were adopted as feature extractor to the MLP, namely, RGB, HSV, HS, YCbCr, CbCr, RGBHSVYCbCr, and RGBHSCbCr, as one can see in Table 1. For the evaluation and comparison of the results obtained in the classification of pixels in skin and non-skin, the following metrics were used: precision or positive predictive value (PPV), sensitivity or true positive rate (TPR), specificity or negative real rate (SPC), and accuracy (ACC). The mean (μ) and standard deviation (σ) are also shown, the best result of each metric is in bold.

Table 1. Testing of neural networks with different color models.

Colorspace		Metrics			
		PPV	TPR	SPC	ACC
RGB	μ	96.54	93.43	96.42	94.87
	σ	0.0015	0.0046	0.0015	0.0025
HSV	μ	96.24	94.06	96.15	95.08
	σ	0.0018	0.0035	0.0018	0.0021
HS	μ	93.29	90.91	93.11	91.98
	σ	0.0028	0.0038	0.0027	0.0028
YCbCr	μ	96.54	93.47	96.42	94.89
	σ	0.0014	0.0039	0.0014	0.0022
CbCr	μ	93.24	88.24	92.84	90.41
	σ	0.0011	0.0005	0.0010	0.0004
RGBHSVYCbCr	μ	**96.74**	**94.93**	**96.68**	**95.79**
	σ	0.0018	0.0038	0.0018	0.0024
RGBHSCbCr	μ	96.68	94.80	96.61	95.69
	σ	0.0026	0.0017	0.0026	0.0019

We note in Table 1, that the approach to despise the illumination channel of HSV and YCbCr color systems was inefficient, since the networks with HSV and YCbCr inputs presented higher results than HS and CbCr networks, respectively, for all metrics, contradicting the results obtained in [Hajiarbabi & Agah 2015]. The network with the colorspace RGBHSVYCbCr as the input showed better accuracy among all other networks, including the RGBHSCbCr, this result is understandable when analyzing separately the results of HSV and YCbCr networks, which proved to be more efficient than

HS and CbCr. The best-performing network was used in Section 5.2 for the RPP-SC experiments.

5.2. RPP-SC experiments

Two different databases were used in the experiments presented in this Section, RetargetMe [Rubinstein et al. 2010] consisting of 80 images and the NRID [Fang et al. 2012] database containing 35 images. Both databases are diversified and widely used for evaluation of image resizing tools and cover images of people and landscapes.

Fig. 4 provides examples that demonstrate execution process the SC and RPP-SC methods, where were applied a reduction of 25% of the total size of image. Fig. 4(a) are the original images; Fig. 4(b) shows the seam selected by the SC method to be removed; Fig. 4(c) shows the images after removal of the seams by SC; Fig. 4(d) shows the seam selected by the RPP-SC method to be removed; Fig. 4(e) shows the images after the removal of the seams by the RPP-SC method.

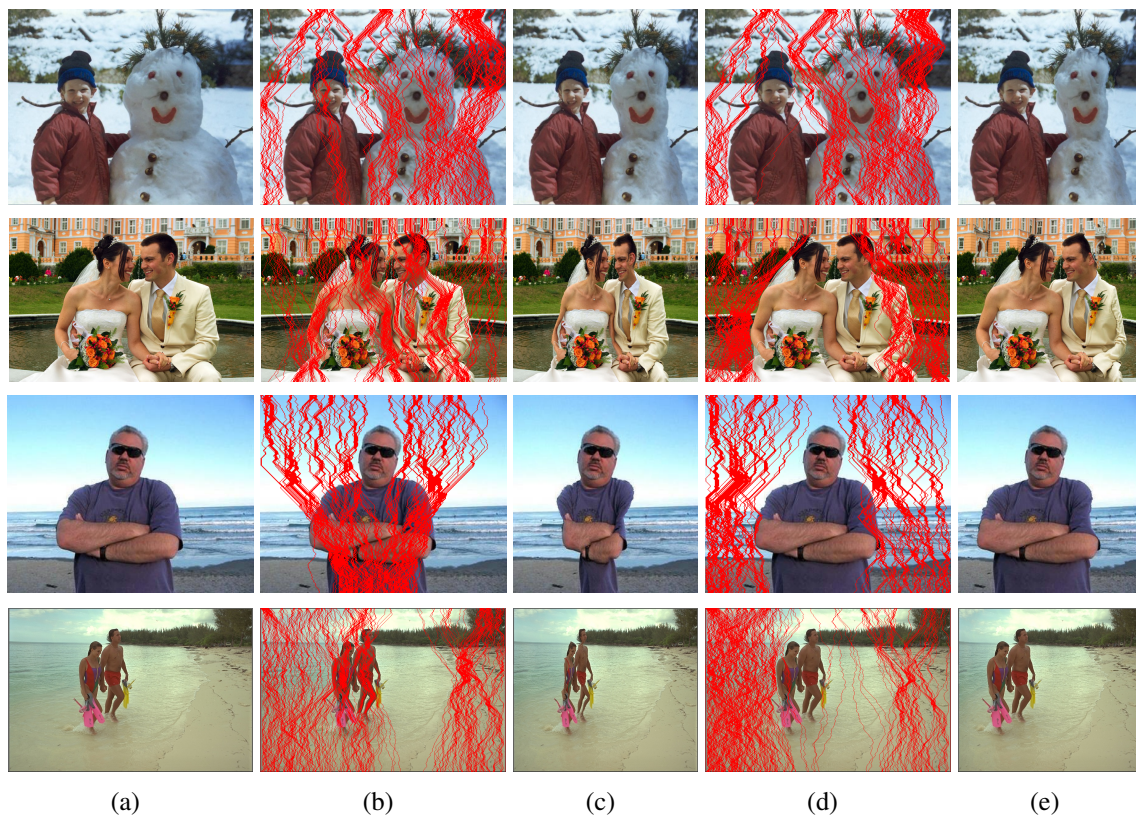

    (a)   (b)   (c)   (d)   (e)

Figure 4. Comparison between the methods with reduction of 25%. (a) Original image. (b) Selection of SC seam. (c) Result of the SC method. (d) Selection of RPP-SC seam. (e) Result of the RPP-SC method.

When analyzing the Fig. 4 we can see that SC method presented distortion of relevant information of the people present in the image. RPP-SC method presented results visibly superior to the original SC, minimizing the number of seams that passing in areas classified as human skin.

6. Validation

In this section, we present a comparison of the proposed method with the method that inspired this article, Seam Carving. For the comparison of approaches, quality analysis metrics were used in resized images. The IR-SSIM and SIFT flow methods were used.

The image retargeting SSIM (IR-SSIM) [Fang et al. 2014] generates a structural similarity map (SSIM) between the reference image and resized image. The reference image is used to generate a (bottom-up) saliency map that is computed using data-driven, task-independent visual attention models, and a top-down saliency map is created based on task-dependent, high-level prior knowledge about the scene. The IR-SSIM algorithm used in these experiments uses a face detector to calculate the top-down overhang map. The final IR-SSIM scoring of the scaled image is calculated by weighting the overall ledge map by the SSIM map.

The SIFT flow [Liu et al. 2011] method is based on optical flow methods that produce dense pixel-to-pixel correspondences between two images. SIFT flow combines optical flow with SIFT descriptors that are extracted from each pixel to characterize local structures of image and encode reference information of the pixel. The SIFT flow metric measures the distortion of the resizing by the distance between corresponding flow vectors between images, the distance is a representation of the visual impact of the resizing. The lower this cost is, the less noticeable the visual impact should be.

The values obtained with IR-SSIM and SIFT flow metrics and the total of seams that pass through areas classified as skin for the SC and RPP-SC methods applied to the NRID and RetargetME databases for 50% reduction is shown in the Table 2. The mean (μ) and standard deviation (σ) of the metrics applied to the database images are used for comparison, the best result of each metric is in bold. For the IR-SSIM metric the higher the score the greater the similarity between the original image and the resized image, while for the SIFT flow metric a lower score represents less content distortion between the reduced image and the original image.

Table 2. Results of the validation experiments.

Database	Metrics		Method	
			SC	RPP-SC
RetargetMe	IR-SSIM	μ	**0.1213**	0.1168
		σ	0.0407	0.0388
	SIFT flow	μ	1.4218×10^5	$\mathbf{1.4153 \times 10^5}$
		σ	1.5829×10^5	1.5825×10^5
	seams	μ	0.7736	**0.4511**
NRID	IR-SSIM	μ	0.1132	**0.1135**
		σ	0.0398	0.0419
	SIFT flow	μ	4.0672×10^4	$\mathbf{4.0607 \times 10^4}$
		σ	3.6965×10^4	3.6976×10^4
	seams	μ	0.7910	**0.4930**

The values obtained with the metrics may not agree with each other because each method uses different ways of inferring how good the result of the resizing was for its original image. Analyzing the results presented in the Table 2 we can see that the RPP-SC method obtained a better result of mean and standard deviation than the SC in most cases, thus showing greater efficiency and reliability of the method. By observing the RPP-SC method applied to the NRID database, the method achieved better values than the SC in all the metrics, this happens because the skin classification for this database achieved visibly better results. We can also observe a considerable reduction in the number of seams that pass in areas classified as skin in the RPP-SC method.

In addition to SC the RPP-SC method was compared with other retargeting operators Nonhomogeneous warping (WARP) [Wolf et al. 2007], Scaling (SCL) [Samadani et al. 2007] and Mutli-operator (MULTIOP) [Rubinstein et al. 2009] the reductions made for this comparison were 25% with dataset NRID as can be seen in Table 3. The mean (μ) and standard deviation (σ) of the metrics applied to the database images are used for comparison, the best result of each metric is in bold.

Table 3. RPP-SC method comparison with other resizing operators.

Metrics		Retargeting operators			
		RPP-SC	WARP	SCL	MULTIOP
IR-SSIM	μ	0.1508	0.1500	0.1493	**0.1580**
	σ	0.0509	0.0484	0.0504	0.0506
SIFT flow	μ	$\mathbf{3.7115 \times 10^3}$	3.8310×10^3	3.8129×10^3	5.2923×10^3
	σ	1.7530×10^3	1.7015×10^3	1.6398×10^3	6.6692×10^3

The results shown in Table 3 show that the RPP-SC method had better results than most operators in most cases. Only for the IR-SSIM metric did the proposed method show a lower result than the MULTIOP operator, a method that brings together a set of operators for scaling. But for the SIFT flow metric the method performed better.

7. Conclusions

We present in this article the RPP-SC method that uses a new energy function for Seam Carving elaborated from a neural network capable of detecting skin color. The network was trained using separately 7 inputs with different color models which were then evaluated taking into account the mean and standard deviation of precision, sensitivity, specificity and accuracy with the objective of finding the best colorspace to classify skin or non-skin pixels skin. Two of the colorspaces tested employed the technique of discarding the illumination channel to avoid problems with the variation of illumination, but were inefficient.

Accuracy and precision have assumed significant values confirming the validity and reliability of the model, so the network can detect the skin correctly in most cases. The results obtained with the quality analysis methods of the resized images show the viability of the proposal, since RPP-SC showed better performance than the results achieved through SC in most cases. The perceptible distortions caused by using the RPP-SC method are minimized compared to those obtained by the SC method. As future work, it

may be suggested to use new input attributes for the neural network, allowing, for example, also to consider region-based and texture-based skin detection methods.

References

Avidan, S. & Shamir, A. (2007). Seam carving for content-aware image resizing. *ACM*, 26(3):10.

Baskan, S., Bulut, M. M., & Atalay, V. (2002). Projection based method for segmentation of human face and its evaluation. *Pattern Recognition Letters*, 23(14):1623–1629.

Bhute, A. N. & Meshram, B. (2014). Content based image indexing and retrieval. *arXiv preprint arXiv:1401.1742*.

Casati, J. P. B., Moraes, D. R., & Rodrigues, E. L. L. (2013). Sfa: A human skin image database based on feret and ar facial images. In *IX workshop de Visao Computational, Rio de Janeiro*.

Dong, W., Zhou, N., Paul, J.-C., & Zhang, X. (2009). Optimized image resizing using seam carving and scaling. In *ACM Transactions on Graphics (TOG)*, volume 28, page 125. ACM.

Fang, Y., Chen, Z., Lin, W., & Lin, C.-W. (2012). Saliency detection in the compressed domain for adaptive image retargeting. *IEEE Transactions on Image Processing*, 21(9):3888–3901.

Fang, Y., Zeng, K., Wang, Z., Lin, W., Fang, Z., & Lin, C.-W. (2014). Objective quality assessment for image retargeting based on structural similarity. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 4(1):95–105.

Habili, N., Lim, C. C., & Moini, A. (2004). Segmentation of the face and hands in sign language video sequences using color and motion cues. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(8):1086–1097.

Hajiarbabi, M. & Agah, A. (2015). Human skin color detection using neural networks. *Journal of Intelligent Systems*, 24(4):425–436.

Han, J., Awad, G., & Sutherland, A. (2009). Automatic skin segmentation and tracking in sign language recognition. *IET Computer Vision*, 3(1):24–35.

Kakumanu, P., Makrogiannis, S., & Bourbakis, N. (2007). A survey of skin-color modeling and detection methods. *Pattern recognition*, 40(3):1106–1122.

Kang, S., Choi, B., & Jo, D. (2016). Faces detection method based on skin color modeling. *Journal of Systems Architecture*, 64:100–109.

Liu, C., Yuen, J., & Torralba, A. (2011). Sift flow: Dense correspondence across scenes and its applications. *IEEE transactions on pattern analysis and machine intelligence*, 33(5):978–994.

Liu, C.-C. & Chung, P.-C. (2011). Objects extraction algorithm of color image using adaptive forecasting filters created automatically. *International Journal of Innovative Computing, Information and Control*, 7(10):5771–5787.

M. Martinez, A. & Benavente, R. (1998). The ar face database. *Tech. Rep. 24 CVC Technical Report*.

Moallem, P., Mousavi, B. S., & Monadjemi, S. A. (2011). A novel fuzzy rule base system for pose independent faces detection. *Applied Soft Computing*, 11(2):1801–1810.

Naji, S., Jalab, H. A., & Kareem, S. A. (2018). A survey on skin detection in colored images. *Artificial Intelligence Review*.

Naji, S. A. (2013). *Human face detection from colour images based on multi-skin models, rule-based geometrical knowledge, and artificial neural network*. PhD thesis, University of Malaya.

Nguyen, D. T., Li, W., & Ogunbona, P. O. (2016). Human detection from images and videos: A survey. *Pattern Recognition*, 51:148–175.

Oliveira, S. A., Bezerra, F. N., & Neto, A. R. R. (2015). Genetic seam carving: A genetic algorithm approach for content-aware image retargeting. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 700–707. Springer.

Perez, M., Avila, S., Moreira, D., Moraes, D., Testoni, V., Valle, E., Goldenstein, S., & Rocha, A. (2017). Video pornography detection through deep learning techniques and motion information. *Neurocomputing*, 230:279–293.

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(5):295–306.

Phung, S. L., Bouzerdoum, A., & Chai, D. (2003). Skin segmentation using color and edge information. In *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings.*, volume 1, pages 525–528. IEEE.

Rubinstein, M., Gutierrez, D., Sorkine, O., & Shamir, A. (2010). A comparative study of image retargeting. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*, 29(6):160:1–160:10.

Rubinstein, M., Shamir, A., & Avidan, S. (2009). Multi-operator media retargeting. *ACM Transactions on graphics (TOG)*, 28(3):23.

Rumelhart, D. E., Hinton, G. E., Williams, R. J., et al. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5(3):1.

Samadani, R., Lim, S. H., & Tretter, D. (2007). Representative image thumbnails for good browsing. In *2007 IEEE International Conference on Image Processing*, volume 2, pages II–193. IEEE.

Senturk, Z. K. & Akgun, D. (2017). Seam carving based image resizing detection using hybrid features. *Tehnicki Vjesnik-Technical Gazette*, 24(6):1825–1833.

Shamir, A. & Avidan, S. (2009). Seam carving for media retargeting. *Communications of the ACM*, 52(1):77–85.

Wolf, L., Guttmann, M., & Cohen-Or, D. (2007). Non-homogeneous content-driven video-retargeting. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–6. IEEE.

Yuetao, D. & Nana, Y. (2011). Research of face detection in color image based on skin color. *Energy Procedia*, 13:9395–9401.

Zhipeng, C., Junda, H., & Wenbin, Z. (2010). Face detection system based on skin color model. In *2010 International Conference on Networking and Digital Society*, volume 2, pages 664–667. IEEE.