IJCTA, 9(37) 2016, pp. 81-88

Image Processing used for Lung Cancer Detection in Medical Images

Neelima Singh* and A. Asuntha**

Abstract: This paper deals with formation of lung cancer detection system through use of various techniques of image processing. Image processing applies mathematical operations for the processing of an image. For this pupose it uses any one form of signal processing with an image as input, such as photograph or video frame. The result of image processing may be either a set of features or parameters associated with the image or an image itself. The system formed accepts any one of medical image within the three choices consisting of MRI, CT and Ultrasound image as input. After preprocessing of image, Canny filter is used for Edge detection. This present work proposes a method to detect the cancerous cells effectively from the CT, MRI scan and Ultrasound images by reducing the detection error made by the physicians' naked eye for medical study based on Superpixel segmentation. Simulation result is obtained for the above discussed system using MATLAB. The result is obtained and comparison is done between all the three images.

Keywords: Superpixel Segmentation, Image processing, Morphological processing, Canny Edge detection, Dilation, Erosion.

1. INTRODUCTION

In present era, Cancer has become one of the highest concern in medical field. Lung cancer is having the smallest survival rate even after diagnosis, which is leading to a gradual increase in the number of deaths every year. The fatality rate from lung cancer is on the peak if compared with all other types of cancer. Survival from lung cancer is directly dependant on the region of it's growth at the detection time. The chances of survival through successful treatment increases if the cancer is successfully without any false negatives during it's earlier stage. Cigarette smoking has caused around 85% lung cancer cases in males and 75% cases in females [1]. Nearly 14.5 million Americans having cancer in history were found to be alive on January 1, 2014. Some of these people were detected recently having lung cancer and are undergoing treatment. About 1,685,210 new cancer cases are expected to be diagnosed in 2016. About 595,690 Americans are speculated to die of cancer in 2016, which is equivalent to an estimate of about 1,630 people per day .Cancer is the second most common cause of death in the US, except heart disease, and accounts for nearly 1 of every 4 deaths. For all types of cancer, the overall survival rate is 63%. Although surgery, radiation therapy and chemotherapy has been used in the treatment of lung cancer, the survival rate for all stages combined was only 14%. Because of tobacco epidemic, the overall cancer death rate grew for the 20th century hiking in 1991 at 215 cancer deaths per 100,000 persons. However, from 1991 to 2012, the rate has dropped 23% because of progress in early detection and treatment.

The basic aim of this paper is to develop an efficient system which is able to detect lung cancer in early stages based on automatic diagnosis of the lung regions included in chest CT, MRI images. This system can also be used for detection of further stages if incase the lung cancer spreads to the other organs through the ultrasound images.

Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) are the types of lung cancer. As non-small cell lung cancer is a more prevalent lung cancer, this project is based on the study and detection

* Dept. of Electronics and Control, SRM University, Chennai, India. *Email: bme.0905601017@gmail.com*
** Dept. of Electronics and Instrumentation, SRM University, Chennai, India. *Email: asuntha.a@ktr.srmuniv.ac.in*

of non-small cell lung cancer. There's a difference between the diagnosis and treatment of non-small cell and small cell lung cancer due to their distinguishable characteristics. There are various ways to detect lung cancer by using Computed Tomography (CT) scan image, Magnetic Resonance Imaging (MRI) scan image, Ultrasound image. Image processing of necessary part of diagnosis is for early diagnosis. For this we develop a system which will help doctors to easily detect cancer for any one of three images given as input and gives proper analysis. CT scan image, MRI scan image and Ultrasound image are the three types of images used. The goal of this study is to design a system which can take any one of the three images and produce an output. The algorithm used here is a powerful algorithm in terms of sensitivity, specificity and accuracy. The proposed model consists of following steps such as: Collect lung CT, MRI and Ultrasound image data set, preprocessing, edge detection and morphological processing. Every step is described in further part.

2. MATERIALS AND METHODS

In the following section we represent the methods used for the proposed model and how preprocessing is being carried out on the images collected. Then we find the best method which can detect cancer and extract cancer tissue information from all the three images. Finally the classification is done amongst all the images .The steps to classification of lung cancer are given in the figure.

The data used for this study includes 5 ultrasound images, 15 MRI and 6 CT scan images of lung. The lungs image have a dimension of 512 × 512.

A. Image pre-processing

Image pre-processing is used for denoising and prepares the images for further steps such as segmentation. It diminishes distortion in image and improves it's important features. Thus we get rectified image. MATLAB software has been used for this purpose.

a) *Image enhancement:* Different image enhancement techniques are used for all the different images This includes smoothing of image also and removal of noises, blurring etc. But in a combined manner, Gabor filter was found to be suitable for all the three images together. The filtering of image proves to be useful for further steps.

b) *Layer separation:* We eliminate the effect of other two colors i.e. red and blue and represent the image in green color. This is to reduce the complexity and proper conversion to gray level.

c) *Gray conversion:* This includes conversion of coloured image with pixels having RGB level into Gray level. A Gray level image can be easily processed in comparison to colored image. The reason is the pixels to be processed separately which have different RGB values. Therefore we go for Gray conversion.

d) *Edge detection:* It is observed that majority of false positives are due to the edge regions [25], which are mainly because of camera defocus. Since edge pixels share similar hue as nodule pixels, they are often mistaken for nodule pixels. Hence, they should be removed prior to cancer nodule detection. The contrast is the difference in visual properties that makes an object distinguishable from other objects and the background.

In a typical medical image, luminance is the most relevant feature between edge pixels and other non-cancerous pixels.

Canny detector [26] is adopted to find edge pixels due to it's fast performance and suitability for real-time detection. Edges are dilated to locate the edge regions using morphological operation, after edge detetion. In detail, the dilation operator needs two inputs as shown in Eq. (1). A is the image to be dilated.

B is a set of coordinate points known as a structuring element (also known as a kernel) which determines the precise effect of the dilation on the input image

$$A \otimes B = \{z \mid B_Z \cap A \neq \varnothing\} \quad (1)$$

This equation is based on obtaining the reflection of B about its origin as well as shifting this reflection by a value z. The dilation of A by B is then the set of all displacements z, such that B and A overlap by at least one pixel. More details can be studied through [27]. Finally, we obtain an image with enlarged edge regions which will be removed by masking.

B. Lung Region Extraction

a) Superpixel Segmentation

With the development of imaging technology, the resolution of sensor becomes higher and higher. First of all the pixels are grouped based on colour and location followed by detection of cancerous nodule at superpixel level which reduces the computational cost and makes the detection of cancer prone area faster.

As a restricted form of region segmentation, superpixels can balance the conflicting goals of reducing image complexity through pixel grouping while avoiding under segmentation [28], [29]. Most superpixel methods often suffer from high computational cost, poor segmentation quality, inconsistent size and shape, or multiple difficult-to-tune parameters. In this section, we introduce an initial seed (cluster center) growth based approach extended from [29] and [30]. To limit the computational load of superpixel segmentation, we made several improvements for speeding up. Our method relies on pixel similarity to find clusters instead of using curve evolution, and thus works much faster than [29]. Our method maintains the simple linear clustering manner as [30] but has some new properties. First, since CT, MRI and Ultrasound images are often noisy because of the poor illumination conditions, we first use Gabor filter to remove noise. Second, cancer prone regions may be irregular or 'spiculated' in a medical image and appear as disjoint segments in further stages. To enforce connectivity, most methods relabel them with the labels of the larger neighboring cluster which may be a normal region. Thus, it is hard to detect cancerous from this superpixel since most pixels are noncancerous. The solution for this problem is to assign a new label to the disjoint segment, and thus the small cancerous region can show up as a single superpixel. Besides, we found that one time segmentation can give satisfactory superpixel results for all the three medical images. So our algorithm is noniterative, therefore can be said as faster than [30]. The details are as follows: (a), (b)

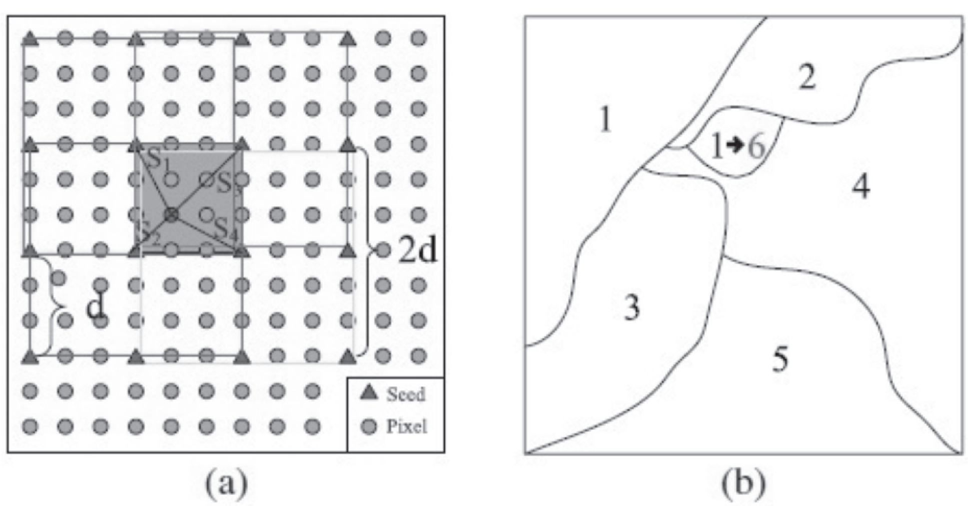

Figure 1: Superpixel segmentation. (a) Similarity calculation. (b) Disjoint superpixel relabel

As most previous methods, we place K seeds in a lattice formation to make the superpixels evenly distributed over the image. The distances between lattice neighbors are all approximately equal to d, where d is $\sqrt{\frac{N}{K}}$,where N is the total number of pixels in an image. The color similarity between pixels can be measured in the CIELab space. It is given as uniform changes of components in CIELab correspond to uniform changes in observed color. The relative affective differences between any two colors in CIELab can be approximated by assuming each color as a point in a 3-D space (with three components: l, a, b) and taking the Euclidean distance between them

$$S_{color} = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \quad (2)$$

Similarity of two pixels is not only related to the color similarity, but also related to the spatial distance as follows:

$$S_{spatial} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

Hence, we define the similarity measure S by combining them together as

$$S = S_{color} + \lambda \times S_{spatial}$$

where S_{color} is the color similarity in the CIELab space and $S_{spatial}$ denotes the spatial distance between two pixels. (x, y) denotes the position of a pixel. The parameter λ is used to balance the contributions of color similarity and spatial distance. As shown in Figure 2(a), the K superpixel seeds are placed with sampling intervals d. Thus, to ensure enough overlapping, the searching space for each seed is set to $2d \times 2d$. Hence, pixels will be covered by its neighboring four seeds, and the similarity between each pixel and the four seeds should be calculated. Finally, each pixel will be associated with the seed that has the closest similarity to it.

However, there may exist a few disjoint segments in the result because the connectivity constraint of the same superpixel is not imposed. This may occur especially for cancerous regions because they are often small in the medical image. Traditional methods will combine the disjoint segments with the neighboring segment to enforce connectivity.

C. Feature Extraction

The Image feature extraction stage is an essential step in our project in image processing techniques which is done by using algorithms and techniques to detect and remove various desired portions or shapes (features) present in an image. Initially segmentation is performed on lung region followed by steps of feature extraction to obtain it's features. At last in accordance with some diagnosis rule the cancer nodules can easily be detected in the lungs. These diagnosis rules can eliminate the false detection of cancer nodules and false positives resulted through segmentation and provides better diagnosis. In the literature we found among the features used in the diagnostic indicators.

- Shape
- Area of interest
- Size of nodule and
- Contrast Enhancement
- Calcification

Similarly, To achieve accurate diagnosis we experimentally found the above suitable texture features As a matter of fact, the first feature (the area of the candidate region or object) is used for:

- Elimination of very small candidate object (Area is less than a thresholding value).
- Elimination of isolated pixels (seen as noise in the segmented image).

The use of the necessary feature usually leads to the elimination of the extra candidate regions that probably will not form a nodule; furthermore it's utilization tends to minimize the computation time required in the upcoming diagnostic steps.

3. RESULTS AND DISCUSSION

In this paper Lung CT, MRI and Ultrasound images used were obtained from a specialist medical imaging center. The image enhancement is done using Gabor filter. After enhancement step, the images were passed from layer separation step and then converted to Gray level image. Canny Edge detection method is used due to it's high performance. For segmentation Superpixel segmentation algorithm was used thus lung region or (ROI) is extracted. The steps applied on Ultrasound, CT and MRI images are shown in Figure 3.1, Figure 3.2 and Figure 3.3 respectively.

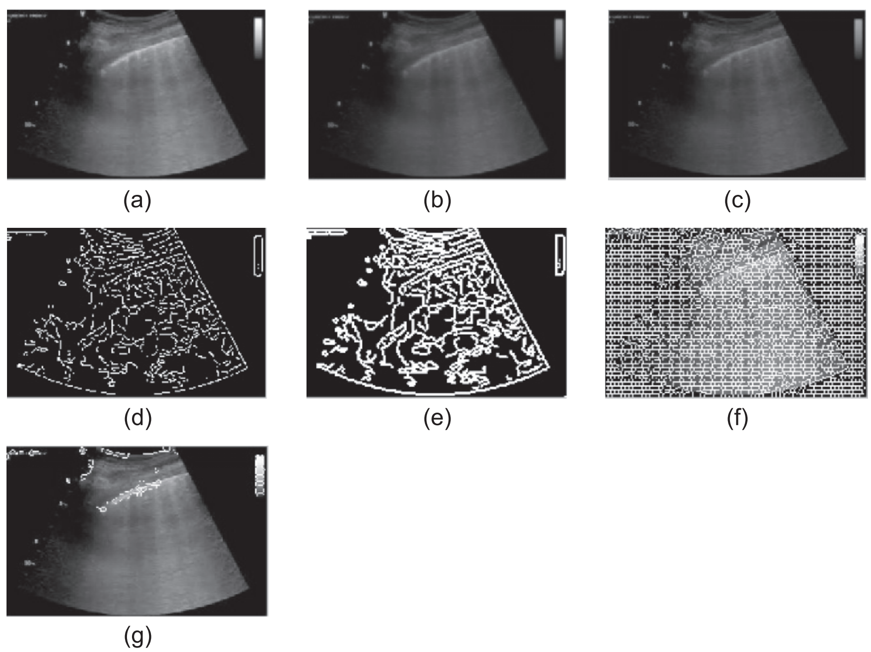

Figure 3.1: Segmentation steps for Ultrasound Image: (a) input denoised image, (b) green layer separation, (c) gray level intensity, (d) edge detection, (e) morphological processing, (f) segmentation, (g) superpixel segmentation

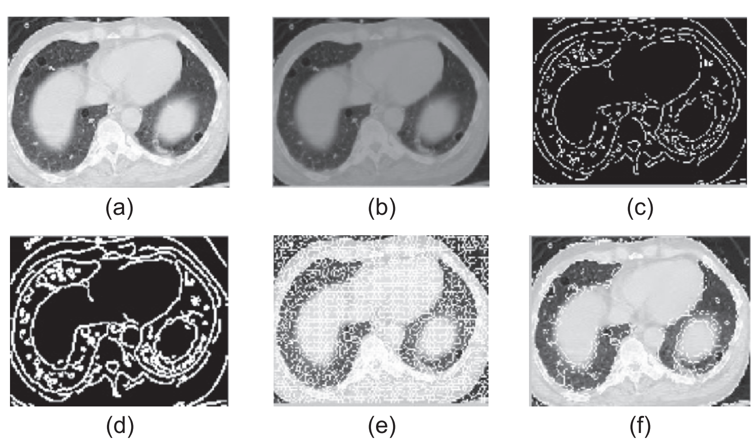

Figure 3.2: Segmentation steps for CT Image: (a) input denoised image, (b) green layer separation, (c) gray level intensity, (d) edge detection, (e) morphological processing, (f) segmentation, (g) superpixel segmentation

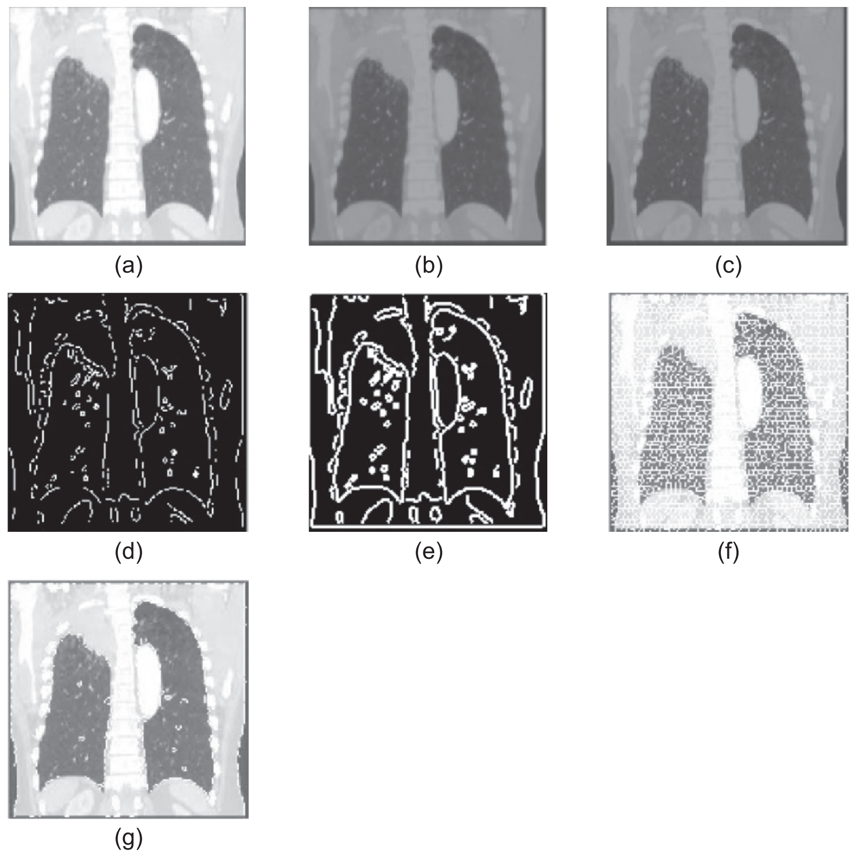

Figure 3.3: Segmentation steps for MRI Image: (a) input denoised image, (b) green layer separation, (c) gray level intensity, (d) edge detection, (e) morphological processing, (f) segmentation, (g) superpixel segmentation

4. CONCLUSION

In this survey of cancer detection in MRI, CT and Ultrasound image, we have studied the major image modalities through image processing. We used canny edge detection method and gabor filter for preprocessing of images. Superpixel segmentation method was used as it gives higher output in comparison to other algorithms in this case due to difference in luminisance of the three images.

5. FUTURE WORK

By the process used complexity is reduced and diagnosis confidence is improved. Canny filter is used for edge detection and finally we go for superpixel segmentation. Further we can do the classification through Pearsons and Spearman algorithm and SVM algorithm and detect the cancer prone region in all the three images.

References

1. Nooshin Hadavi , Jan Nordin, Ali Shojaeipour, "*Lung Cancer Diagnosis using CT-Scan images based on Cellular Learning Automata*", IEEE, 2014.
2. B.V. Ginneken, B. M. Romeny and M.A. Viergever, "*Computer-aided diagnosis in chest radiography: a survey*", IEEE, transactions on medical imaging, Vol. 20, No. 12, (2001).
3. American Cancer Society, "*Cancer Statistics, 2005*", CA: A Cancer Journal for Clinicians, 55: 10-30, 2005,http://caonline.amcancersoc.org/cgi/content/full/55/1/1
4. D. Lin and C. Yan, "*Lung nodules identification rules extraction with neural fuzzy network*", IEEE, Neural Information Processing, Vol. 4, (2002).

5. K. Kanazawa, Y. Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. Kakinuma, M. Kaneko, N. Moriyama and K. Eguchi, "*Computer-aided diagnosis for pulmonary nodules based on helicalCT images*", Compute. Med. Image Graph, Vol. 22, No. 2 (1998), pp. 157-167.

6. A. El-Baz, A. A. Farag, PH.D., R. Falk, M.D. and R. L. Rocco, M.D.,"*detection, visualization, and identification of lung abnormalities in chest spiral CT scans: phase I*", Information Conference on Biomedical Engineering, Egypt(2002).

7. B. Zhao, G. Gamsu, M. S. Ginsberg, L. Jiang and L. H. Schwartz, "*Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm*", journal of applied clinical medical physics, Vol. 4, (2003).

8. The DICOM Standards Committee. DICOM homepage http://medical.nema.org/, September 2004.

9. Linda G. Shapiro and G.C. Stockman., Computer Vision: Theory and Applications. 2001: Prentice Hall

10. B. Magesh, P. Vijaylakshmi, M. Abhiram, "*Computer aided Diagnosis System for identification and classification of Lessions in Lungs*", International Journal of Computer Trends and Technology-May to June Issue 2011.

11. Rachid Sammouda, Jamal Abu Hassan1, Mohamed Sammouda, Abdulridha Al-Zuhairy, Hatem abou ElAbbas, "*Computer Aided Diagnosis System for Early Detection of Lung Cancer Using Chest Computer Tomography Images*", GVIP 05 Conference, 19-21, December 2005, CICC, Cairo, Egypt.

12. Proceedings of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry, HEALTHCOM 2005. pp: 378 –383, 2005.

13. Disha Sharma, Gagandeep Jindal, "*Identifying Lung Cancer Using Image Processing Techniques*", International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011), Vol. 17, pp: 872-880, 2011

14. Ilya Levner, Hong Zhangm,"*Classification driven Watershed segmentation*", IEEE Transactions on Image Processing Vol. 16, No. 5, 2007

15. Anita chaudhary, Sonit Sukhraj Singh, "*Lung Cancer Detection on CT Images Using Image Processing*", International transaction on Computing Sciences, Vol. 4, 2012

16. B.V. Ginneken, B. M. Romeny and M. A. Viergever, "*Computer-aided diagnosis in chest radiography: a survey*", IEEE, transactions on medical imaging, Vol. 20, No. 12, DEC-2001

17. Weiner Filtering, " http://www.owlnet.rice.edu". International Journal of Computer and Electrical Engineering, Vol. 3, No. 5, October 2011

18. Antony Judice A., K. Parimala Geetha, A Novel Assesment of Various Bio-Imaging Methods for Lung Tumour Detection and Treatment by 4-D and 2-D CT Images, pp 1-8, 2013.

19. Vijay A.Gajdhane, Deshpande L.M. Detection of Lung Cancer Nodule on Computed Tomography Images by Using Image Processing, Volume 16, Issue 5, pp 28-35, (Sep – Oct. 2014).

20. Tao Suna, Jingjing Wanga, Xia Li Pingxin Lvc, Fen Liua, Yanxia Luoa,Qi Gaoa, Huiping Zhua,Xiuhua Guo, Comparative evaluation of support vector machines for computer aided diagnosis of lung cancer in CT based on a multi-dimensional data set, pp 519-524, 2013.

21. B. Penna, T. Tillo, M. Grangetto, E. Magli, and G. Olmo, "*A technique for blood detection in wireless capsule endoscopy images*," in Proc. 17th Eur. Signal Process. Conf., Glasgow, Scotland, 2009, pp. 1864–1868.

22. Y. Jung, Y. Kim, D. Lee, and J. Kim, "*Active blood detection in a high resolution capsule endoscopy using color spectrum transformation*," inProc. Int. Conf. Biomed. Eng. Inf., May 2008, Vol. 1, pp. 859–862.

23. P. Lau and P. Correia, "*Detection of bleeding patterns in wce video using multiple features*," in Proc. IEEE 29th Annu. Int. Conf. Eng. Med. Biol.Soc., Aug. 2007, pp. 5601–5604.

24. B. Li and M. Meng, "*Computer-aided detection of bleeding regions for capsule endoscopy images*," IEEE Trans. Biomed. Eng., Vol. 56, No. 4, pp. 1032–1039, Apr. 2009.Y. Fu, M. Mandal, and G. Guo,

25. "*Bleeding region detection in wce images based on color features and neural network*," in Proc. IEEE 54th Int. Midwest Symp. Circuits Syst., Aug. 2011, pp. 1–4.

26. J. Canny, "*A computational approach to edge detection*," IEEE Trans. Patt. Anal. Mach. Intell., Vol. 8, No. 6, pp. 679–698, Nov. 1986.

27. R. C. Gonzalez and R. E.Woods, Digital Image Processing, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2006.

28. X. Ren and J. Malik, "*Learning a classification model for segmentation*," in Proc. IEEE 9th Int. Conf. Comput. Vis., Oct. 2003, Vol. 1, pp. 10–17.

29. A. Levinshtein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson, and K. Siddiqi, "*Turbopixels: Fast superpixels using geometric flows*," IEEE Trans. Patt. Anal.Mach. Intell., Vol. 31, No. 12, pp. 2290–2297, Dec. 2009.
30. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk,"*Slic superpixels compared to state-of-the-art superpixel methods*," IEEE Trans. Patt. Anal.Mach. Intell., Vol. 34, No. 11, pp. 2274–2282, Nov. 2012.
31. C.-C. Chang and C.-J. Lin, "*LIBSVM: A library for support vector machines*,"ACM Trans. Intell. Syst. Technol., Vol. 2, pp. 27:1–27:27, 2011.
32. J. Berens, G. Finlayson, and G. Qiu, "*Image indexing using compressed colour histograms*," IEE Proc. Vis. Imag. Signal Process., Vol. 147, No. 4, pp. 349–355, Aug. 2000.