Performance Measure Based Segmentation Techniques for Skin Cancer Detection

Ginni Arora[1], Ashwani Kumar Dubey[2(✉)], and Zainul Abdin Jaffery[3(✉)]

[1] Amity Institute of Information Technology, Amity University Uttar Pradesh, Noida, UP, India
garora@amity.edu
[2] Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India
dubey1ak@gmail.com
[3] Department of Electrical Engineering, Jamia Millia Islamia, New Delhi, India
zjaffery@jmi.ac.in

Abstract. Skin Cancer is a very common form of cancer which initially starts with investigation and analysis going through biopsy and examination. Doing analysis is the most challenging task as it depends on appearance of skin lesion. Computer Aided Diagnostic (CAD) system have been developed for skin cancer detection which goes through various phases starting from pre-processing, segmentation, feature extraction & selection and classification of cancer type. Segmentation is an important as well as difficult phase which extracts the lesion from non-lesion area depending on variation in terms of color, texture, size and shape. In this paper, different segmentation techniques have been discussed, Otsu thresholding as Pixel Based Segmentation, Canny edge detection as Edge based Segmentation, Watershed as Region Based Segmentation and K-Means as Clustering based Segmentation. The performance of techniques have been measured by Peak Signal Noise Ratio (PSNR), Mean Square Error (MSE) and Structure Similarity Index Measure (SSIM) using MATLAB.

**Keywords:** MSE · PSNR · Skin cancer · Segmentation · SSIM

1 Introduction

Skin cancer is a type of cancer which is more common that usually begins within the skin cell. This type of skin disease (non-melanoma) is usually started in the basal cells or squamous cells. These cells are found at the base of the outer layer of the skin. The exposure of the skin to sunlight is considered the main cause of the development of the most basal and squamous cell cancers. Dermoscopy is a predominantly and a very helpful standard method to diagnose the malignancy [1].

An expected 87,110 new instances of melanoma will be analyzed in the US in 2017. Frequency rates are higher in ladies than in men before age 50, however by age 65, rates in men are twofold those in ladies, and by age 80 they are triple. This example reflects age and sex contrasts in word related and recreational presentation to bright radiation (counting the utilization of indoor tanning), and maybe early discovery practices and utilization of human services. In 2017, an expected 9,730 passing's from melanoma will

happen. Despite the fact that the US Preventive Services Task Force as of late announced that there is deficient proof to prescribe for or against visual skin examination by a clinician for individuals at normal danger of skin tumor, the most ideal approach to distinguish skin malignancy early is to know about new or changing skin developments, especially those that look abnormal. Any new sores, or a dynamic change in an injury's appearance (size, shape, or shading, and so on), ought to be assessed instantly by a doctor. The ABCDE run plots cautioning indications of the most well-known kind of melanoma: A is for asymmetry (one portion of the mole does not coordinate the other half); B is for fringe anomaly (the edges are battered, indented, or obscured); C is for shading (the pigmentation is not uniform, with variable degrees of tan, darker, or dark); D is for distance across more noteworthy than 6 mm (about the measure of a pencil eraser); and E is for advancement, which means an adjustment in the mole's appearance [2].

Initially skin cancer starts with investigation and analysis, a biopsy and examination. Analysis through images is a challenging task as appearance of skin lesion can make the difference. CAD (Computer Aided Diagnostic System) is built for automated analysis of skin lesion but accuracy is still a demanding factor. These systems constitutes of following phases: 1. Pre-processing, 2. Segmentation, 3. Feature Extraction, and 4. Classification. Segmentation is an important as well as difficult phase which extracts the lesion from non-lesion area depending on variation in terms of color, texture, size and shape. Figure 1 shows sample images of different skin cancer types from ISIC (International Skin Imaging Collaboration) Archive [3]. These images are further used for
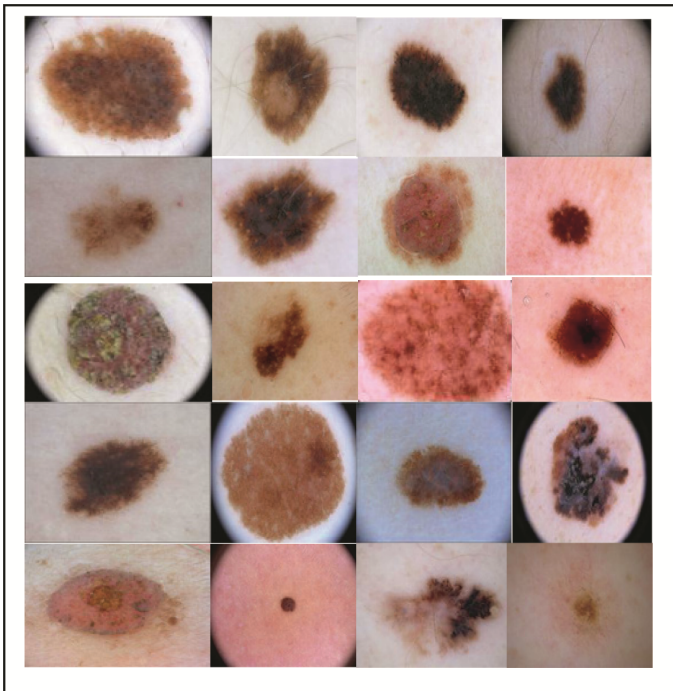
Fig. 1. Skin cancer images

segmentation process. These segmentation techniques broadly classified as: Pixel Based, Edge Based, Region Based and Clustering Based segmentation. In this paper focus has been given on standard techniques defined under broad categories.

2   Segmentation Techniques

Image segmentation is a division process in which digital images are divided into multiple segments in such a way that it is more meaningful as well as easier to evaluate. It is particularly used in finding the object and boundary of an image which is processed by allotting a tag to each and every pixel of the image, thus pixels having similar tag shares a certain characteristic. Otsu thresholding is Pixel based segmentation, Canny edge detection is Edge based Segmentation, Watershed as Region Based Segmentation and K-Means as Clustering based Segmentation. Each of these techniques has different methods of segmentation.

Otsu Thresholding skin lesion image, on the basis of gray level distribution separate the objects of interest from the background selection of an optimal gray level threshold value have become a basic objective. In thresholding, binary images is created based on grey level by converting pixels having below threshold to zero whereas pixel having about that threshold to one. Otsu's thresholding chooses the threshold in order to minimise the variation in intraclass of white and black pixels of threshold [4].

Canny edge detector is defined as the standard benchmark for comparing with other edge detection methods.

Canny detector algorithm consists of four steps as mentioned: (1) usage of Gaussian convolution to blur the image and reduce noise effects (2) usage of Sobel operator and finding the 2D spatial gradient to determine the edge directions and strength (3) use edge direction to search and suppress the non-edge pixels (4) using two high and low thresholds, eliminate edges which are broken, pixels which crosses the high threshold and also lies in between the high and low thresholds that are adjacent to the edge are edge pixels. However Canny edge detector can detect pixels which is near to the edge, but its accuracy is low in finding the exact edges [5].

Watershed segmentation algorithm is considered as half breed technique between group based segmentation and edge based segmentation. The objective of the watershed segmentation is to find the watershed line in the image provided in order to find prominent regions. The concept behind this algorithm is based on flooding water process in topographic surface, flowing from lowest basin to highest peak, water flows into the holes and flood the surface. Thus dam is built to prevent merging of two different basins in which boundary of two connected dam is considered as watershed edges. Though the resultant segmented image leads to the over divided of regions and the time consumption is high, it has advantage over continuity of region's boundaries [4].

As compared to other segmentation techniques, k-means clustering differs from other types of segmentation techniques that the desired numbers of segment or cluster were specified by the analyst. During the classification procedure in K-means, the distances of cluster centers from each respondent are calculated. This procedure is repeated until the distance between skin lesion cluster centers is maximized while respondents are

assigned to the cluster with the nearest Centre. Conversion of RGB to any other color requires number of clusters and sensitive to noise. The K-means technique is highly preferred over other techniques because it is simple, easy and versatile in its implementation. And, it is also computationally faster than hierarchical clustering with linear time and storage complexity. However, one of the most disadvantages of this technique is that it requires number of clusters and is sensitive to noise [5].

3   Performance Measures

There are many types of performance measurement available to evaluate on the performance analysis of different methods in image segmentation. Among those methods, the most efficient technique is selected with the help of performance analysis. This analyses thoroughly on the parameters value which can be either low or high with respective to the type of parameter used.

List of quality parameters used for image segmentation are as follow:

- PSNR (Peak Signal to noise)
- MSE (Mean Square Error)
- SSIM (Structural Similarity Index)

3.1   PSNR

It is a measurement of quality of image in mathematical form which is based on pixel difference between two images and is measured in terms of decibels scale [6]. PSNR value of image is directly proportional to the quality of an image i.e. higher the value betters the quality. It is usually used as a measure to reconstruct the quality of an image. In case, if, the value of PSNR is above 30, clarity in output image is 100%. PSNR of two images is calculated as given by [7] shown in (1):

$$Y_{PSNR} = 10 \log_{10}(\frac{MAX}{X_{MSE}}) \tag{1}$$

Where,

MSE   = Mean Square error
MAX   = value of maximum pixel in the image, usually it's 255 for 8 bit image

3.2   MSE (Mean Square Error)

It is calculated by every picture element in which squared differences of all the pixels are added and divides it by total number of pixel count. In segmented image MSE can be calculated by using the equation given by [8] shown in (2):

$$MSE = \frac{1}{MN} \sum_{y=1}^{M} \sum_{x=1}^{N} \left[I(x, y) - I'(x, y) \right]^2 \tag{2}$$

Where, M and N are number of rows and columns in the given image. I(x, y) is the original image, I'(x, y) is the approximated version. Unlike PSNR, in MSE lower the value leads to the less error obtain higher quality of segmented image.

3.3   SSIM (Structural Similarity Index)

It is a strategy for measuring the similarity between two images based on its structure, contrast and luminance from the same image captured. The SSIM measures the image quality by referencing the uncompressed image so it is also called as reference metric. SSIM is calculated using the following equation given by [9] shown in (3):

$$SSIM(x, y) = \left(2\mu_x\mu_y + c1 \right)(2\sigma_{xy} + c2) / \left(\mu_{x+}^2\mu_{y+}^2 c1 \right)(\sigma_{x+}^2\sigma_y^2 + c2) \tag{3}$$

Where μ_x is the average of x, μ_y is the average of y, σ_{xy} is covariance of x and y, σ_y^2 is the variance of y and σ_X^2 is the variance of x.

4   Experimentation and Results

Based on the selected 10 images from Fig. 1, performance measures are calculated for each of the different type of segmentation techniques. As discussed above the property of each performance technique, the value of mean square error should be less, while the higher the value of PSNR, the better it is. The MSE and the PSNR mainly estimate the absolute error while the SSIM takes into consideration the image degradation as perceived change in structural information. The value of SSIM should be lower. In the above calculations, canny segmentation technique produces the desired result. The MSE value is the less; PSNR is high while the SSIM is also less as desired. This means canny gives a closer view of segmented skin cancer image. Table 1 shows different segmentation techniques with their MSE, PSNR and SSIM values. Lowest MSE, higher PSNR and closest to 1 SSIM is highlighted in bold for each segmentation. Figure 2 represents the graphical view of segmentation techniques based on performance measures. Lowest MSE with redline, higher PSNR with blue line and SSIM nearest to 1 with green line is accurate through canny segmentation technique.

Table 1. Segmentation techniques with performance measures

| Segmentation technique	Image sample no	MSE	PSNR	SSIM
Threshold	S1	255	24.0994	1
	S2	254.84	24.1021	1
	S3	255	24.0994	1
	S4	255	24.0994	1
	S5	255	24.0994	1
	S6	255.02	24.09913	1
	S7	255.03	24.09886	1
	S8	255.047	24.09859	1
	S9	255.06	24.09832	1
	S10	255.078	24.09805	1
Canny	S1	0.5816	50.5187	0.3583
	S2	0.5569	50.7069	0.3738
	S3	0.9309	48.4759	0.0484
	S4	0.9315	48.4728	0.0354
	S5	0.7749	49.2726	0.1828
	S6	0.8052	49.1055	0.156
	S7	0.6349	50.1379	0.3109
	S8	0.9413	48.4277	0.0359
	S9	0.8398	48.9233	0.1223
	S10	0.7832	49.226	0.1607
Watershed	S1	30.8132	33.2774	0.9491
	S2	73.7266	29.4886	0.9212
	S3	70.0962	29.7079	0.846
	S4	63.698	30.1235	0.8451
	S5	90.4841	28.5991	0.901
	S6	117.498	27.4645	0.8147
	S7	118.3323	27.4338	0.8809
	S8	147.3902	26.4801	0.7437
	S9	60.6543	30.3362	0.9196
	S10	91.4905	28.551	0.8001
K-means	S1	234.2328	24.4683	0.0195
	S2	188.9799	25.4006	0.1673
	S3	167.9884	25.912	0.2232
	S4	241.1214	24.3424	0.0403
	S5	215.7924	24.8242	0.0767
	S6	220.0032	24.7405	0.0677
	S7	213.1593	24.8778	0.1145
	S8	171.3275	25.8265	0.2652
	S9	209.9161	24.9443	0.1445
	S10	220.1154	24.7383	0.0619
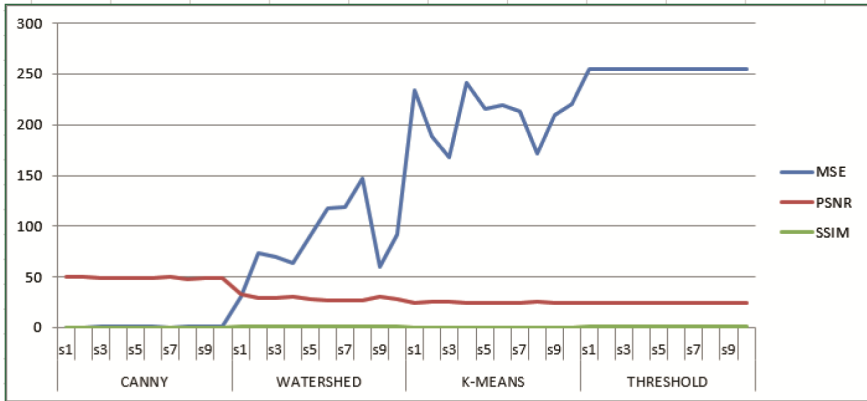
Fig. 2. Analysis of segmentation techniques (Color figure online)

5   Conclusion

Skin cancer cases are increasing day by day so early detection can reduce the chances to avoid deadliest stage. Computer Aided Diagnostic systems are used for this purpose in which segmentation process plays a major role for further classification. This paper highlights major segmentation techniques for differentiating skin lesion area. Further, analyses of these techniques are done through the Mean Square Error, Peak Signal Noise Ratio and Structure Similarity Index performance measures. Canny Segmentation technique resulted to be the best for segmentation of skin cancer images due to low MSE value, high PSNR value and SSIM nearest to 1.

References

1. Urooj, S., Singh, S.: A novel computer assisted approach for diagnosis of skin disease. In: International Conference on Computing for Sustainable Global Development, pp. 1585–1590 (2015)
2. Cancer Facts and Figures. (2017). American Cancer Society. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2017.html. Accessed 31 July 2017
3. https://isic-archive.com/. Accessed 31 July 2017
4. Farooq, M.A., Azhar, M.M., Raza, R.H.: Automatic Lesion Detection System (ALDS) for skin cancer classification using SVM and neural classifiers. In: International Conference on Bioinformatics and Bioengineering, pp. 302–308 (2016)
5. Revathi, V.L., Chithra, A.S.: A review on segmentation techniques in skin lesion images. Int. Res. J. Eng. Technol. **2**(9), 2598–2603 (2015)
6. Iqbal, S., Sheetlani, J.: Application of modified K means clustering algorithm in segmentation of medical images of brain tumor. Biosci., Biotechnol. Res. Asia **14**, 735–739 (2017)

7. Jain, S., Goswami, S.: An efficient approach of segmentation and blind deconvolution in image restoration. IOSR J. Comput. Eng. **17**(6), 41–46 (2015)
8. Wang, Z., Bovik, A.C.: Mean squared error: love it or leave it? IEEE Sig. Process. Mag. 99–117 (2009)
9. Gobindchandra, K., Kumar, S.: Analysis of image segmentation techniques. Int. Res. J. Comput. Sci. **2**(6), 45–53 (2015)