International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

Skin Lesion Image Segmentation Based on C-Means Clustering Algorithm

Deepak Kourav, Abhinav Kathal

Abstract*: *A skin lesion is an abnormal lump, bump, and ulcer, sore or colored area on the skin. There are many types of skin segmentation. In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. In this ipaper using C-Means Clustering approach for skin lesion image segmentation so that detection and recognition of skin decease will easy to understand by patient and biomedical industries.
Keywords*: *Skin. Lesion. Segmentation, C-means, cluster.

I. INTRODUCTION

Skin lesion is commonly located on the face, neck, shoulders, chest, and upper back. Breakouts on the skin composed of blackheads, whiteheads, pimples, or deep, painful cysts and nodule. May leave scars or darken the skin if untreated, Red, painful, fluid-filled blister that appears near the mouth and lips. Affected area will often tingle or burn before the sore is visible. Outbreaks may also be accompanied by mild, flu-like symptoms such as low fever, body aches, and swollen lymph nodes.

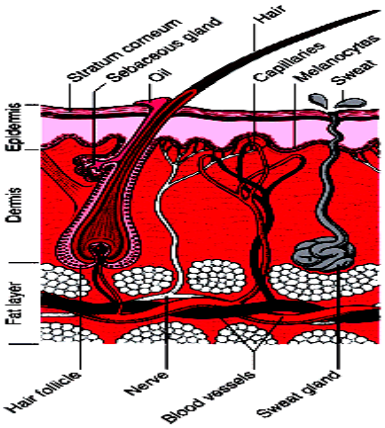

Figure 1: Cross section of skin and skin structure

Skin lesion or cancer image segmentation partitions a skin lesion image into distinct regions containing each pixels with similar attributes. To be meaningful and useful for image analysis and interpretation, the regions should strongly relate to depicted objects or features of interest. Meaningful Segmentation is the first step from low-level image processing transforming a grey scale or color image into one or more other images to high-level image description in terms of features, objects, and scenes. The success of image analysis depends on reliability of segmentation, but an accurate partitioning of an image is generally a very challenging problem.

Segmentation technique Description Advantages Disadvantages Thresholding Method based on the histogram peaks of the image to find particular threshold values no need of previous information, simplest method highly dependent on peaks, spatial details are not considered Edge Based Method based on discontinuity detection good for images having better contrast between objects not suitable for wrong detected or too many edges Region Based Method based on partitioning image into homogeneous regions more immune to noise, useful when it is easy to define similarity criteria expensive method in terms of time and memory Clustering Method based on division into homogeneous clusters fuzzy uses partial membership therefore more useful for real problems determining membership function is not easy Watershed Method based on topological interpretation results are more stable, detected boundaries are continuous complex calculation of gradients PDE Based Method based on the working of differential equations fastest method, best for time critical applications more computational complexity ANN Based Method based on the simulation of learning process for decision making no need to write complex programs more wastage of time in training. Segmentation can be considered the first step and key issue in object recognition, scene understanding and image understanding. Applications range from industrial quality control to medicine, robot navigation, geophysical exploration, and military applications. In all these areas, the quality of the final result depends largely on the quality of the segmentation Clustering is a process whereby a data set is replaced by clusters, which are collections of data points that "belong together". It is natural to think of image segmentation as clustering, grouping those pixels that have the same colour and/or the same texture. Clustering methods can be divided into two basic types: hierarchical and partitional clustering. Within each of the types there exists a wealth. Hierarchical clustering proceeds successively by either merging smaller clusters into larger ones (agglomerative algorithms), or by splitting larger clusters (divisive algorithms).

Revised Manuscript Received on October 30, 2019.
***** Correspondence Author
Dr. Deepak Kourav *, Department of Electronics and Communication Engineering, NRI Institute of Research and Technology, Bhopal, India. Email: deepkourav@gmail.com
Abhinav Kathal, Department of Electronics and Communication Engineering, NRI Institute of Research and Technology, Bhopal, India.

The clustering methods differ in the rule by which it is decided which two small clusters are merged or which large cluster is split. The final result of the algorithm is a tree of clusters called a dendogram, which shows how the clusters are related. By cutting the dendogram at a desired level, a clustering of the data items into disjoint groups is obtained. On the other hand, partitional clustering attempts to directly decompose the data set into a set of disjoint clusters. An objective function expresses how good a representation is, and then the clustering algorithm tries to minimize this function in order to obtain the best representation. The criterion function may emphasize the local structure of the data, as by assigning clusters to peaks in the probability density function, or the global structure. Typically the global criteria involves minimizing a measure of dissimilarity for the samples within each cluster, while maximizing the dissimilarity between different clusters. The most commonly used partitional clustering method is the K-means algorithm, in which the criterion function is the squared distance of the data items from their nearest cluster centroids. Clustering methods, even as thresholding methods, are global and do not retain positional information. The major drawback of this is that it is invariant to spatial rearrangement of the pixels, which is an important aspect of what is meant by segmentation. Resulting segments are not connected and can be widely scattered. Some attempts have been made to introduce such information using pixels coordinates as features. However, this approach tends to result in large regions being broken up and the results so far are no better than those that do not use spatial information. The need to incorporate some form of spatial information into the segmentation process, led to the development of methods where pixels are classified using their context or neighborhood.

II. PROBLEM FORMULATION

From the review it can be conclude that the main issue with skin lesion, no efficient approach to achieve necessary parameters of such images. Mostly k means cluster method is used for such type of work and it calculated only one or two parameters. It is also observed that overall accuracy is not more than 80%.

III. CLUSTERING APPROACH

They are different types of clustering methods, including-

- Partitioning methods.
- Hierarchical clustering.
- Fuzzy clustering.
- Density-based clustering.
- Model-based clustering.

Cluster analysis itself is not one specific algorithm, but the general task to be solved. It can be achieved by various algorithms that differ significantly in their understanding of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with small distances between cluster members, dense areas of the data space, intervals or particular statistical distributions. Clustering can therefore be formulated as a multi-objective optimization problem. The appropriate clustering algorithm and parameter settings (including parameters such as the distance function to use, a density threshold or the number of expected clusters) depend on the individual data set and intended use of the results. Cluster analysis as such is not an automatic task, but an iterative process of knowledge discovery or interactive multi-objective optimization that involves trial and failure. It is often necessary to modify data preprocessing and model parameters until the result achieves the desired properties.

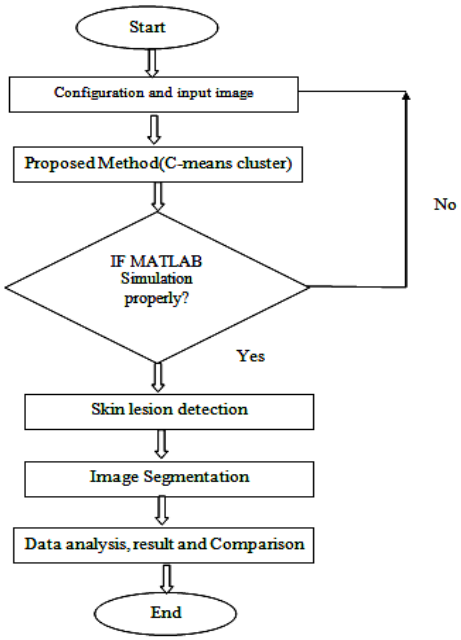

Figure 2: Flow Chart

Fuzzy c-means Clustering Method

- Let $X=\{x1, x2,\ldots, xn\}$ be a set of given data. A fuzzy pseudopartition or fuzzy c-partition of X is a family of fuzzy subsets of X, denoted by $P=\{A1, A2,\ldots, Ac\}$, which satisfies

for all $k\in Nn$ and

for all $i\in Nc$, where c is a positive integer.

- Ex: Given $X=\{x1, x2, x3\}$ and $A1=0.6/x1+1/x2+0.1/x3$

$A2=0.4/x1+0/x2+0.9/x3$, then

$\{A1, A2\}$ is a fuzzy pseudopartition or fuzzy 2-partition of X.

- Given a set of given data $\boldsymbol{X}=\{\boldsymbol{x1}, \boldsymbol{x2},\ldots, \boldsymbol{xn}\}$, where $\boldsymbol{xk}$, in general, is a vector

$\boldsymbol{xk}=\{xk1, xk2,\ldots, xkn\}\in Rp$

for all $k\in Nn$, the problem of fuzzy clustering is to find a fuzzy pseudopartition and the associated cluster centers by which the structure of the data is represented as best as possible. This requires some criterion expressing the general idea that associations be strong within clusters and weak between clusters.

- Given a pseudopartition *P*={A1, A2,…, Ac}, the c clusters, ***v1***, ***v2***,…, ***vn*** associated with the partition are calculated by the formula

$$v_i = \frac{\sum_{k=1}^{n} [A_i(x_k)]^m x_k}{\sum_{k=1}^{n} [A_i(x_k)]^m}$$

for all $i \in N$ number that governs the influence o

- Observe that the vector ***vi*** calculated above, which is viewed as the cluster center of the fuzzy class *Ai*, is actually the weighted average of data in *Ai*.

- The weight of a datum ***xk*** is the *m*th power of the membership grade of ***xk*** in the fuzzy set *Ai*.

where ||·|| is some inner product-induced norm in space RP and || ***xk- vi***||2 i.

- This $J_m(P) = \sum_{k=1}^{n} \sum_{i=1}^{c} [A_i(x_k)]^m \| x_k - v_i \|^2$ 1 sum of distances between cluster centers and elements in the corresponding fuzzy clusters.

- The algorithm is based on the assumption that the desired number of clusters *c* is given and, in addtion, a particular distance, a real number $m \in (1,\infty)$, and a small positive number ε, serving as a stopping criterion, are chosen.

- Step 1: Let t=0. Select an initial fuzzy pseudopartition *P(0)*.

- Step 2: Calculate the *c* cluster centers for *P(t)* and the chosen value of *m*.

- Step 3: Update *P(t+1)* by the following procedure. For each $xk \in X$, if

|| ***xk- vi***||2>0 for all $i \in Nc$, then define

$$A_i^{(t+1)}(x_k) = [\sum_{j=1}^{c} (\frac{\| x_k - v_i^{(t)} \|^2}{\| x_k - v_j^{(t)} \|^2})^{\frac{1}{m-1}}]^{-1}$$

Figure 3: Input images of skin lesion

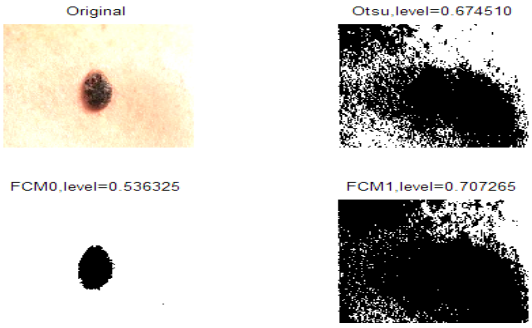

Figure 4: Segmentation of image1(a) Original image (b) Otsu approach (c) FCM(Fuzzy C-means) (d) FCM1

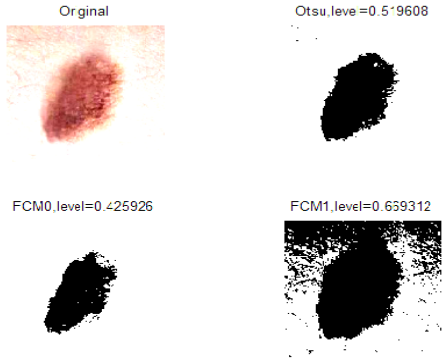

Figure 5: Segmentation of image2(a) Original image (b) Otsu approach (c) FCM(Fuzzy C-means) (d) FCM1

Table -1 Simulation Result summery

Sr. no.	Test Image	OTSU	FCM1	Fuzzy C-Means
1	Image1	0.6745	0.7072	0.5363
2	Image2	0.5196	0.6693	0.4259
3	Image3	0.6666	0.7588	0.4725
4	Image4	0.5372	0.6529	0.3823
5	Image5	0.5803	0.8292	0.5288
6	Image6	0.4686	0.6882	0.3410
7	Image8	0.5372	0.6529	0.3823
8	Image9	0.6666	0.7588	0.4725
9	Image10	0.7746	0.8447	0.7036
10	Image11	0.5725	0.6674	0.4627

According to result of table 1 and segmented figures, it is clearing showing fuzzy C-means clustering approach give better result rather than other approaches.

IV. CONCLUSION

This paper proposed an algorithm for detecting and segmenting skin lesion images. A model for human skin lesion distribution is built using a database of labeled skin pixels. A clustering algorithm based on fuzzy C-means has been trained using the same database of skin pixels. An automatic constructive algorithm has been followed to determine the optimal number of different components required to fit the given data. A skin model, along with a FCM for skin background pixels, is used to compute the probability of every pixel in an input color image to represent skin.

REFERENCES

1. A. Agarwal, A. Issac, M. K. Dutta, K. Riha and V. Uher, "Automated skin lesion segmentation using K-Means clustering from digital dermoscopic images," 2017 40th International Conference on Telecommunications and Signal Processing (TSP), Barcelona, 2017, pp. 743-748.
2. H. Ozkan, R. Gurleyen, E. Usta and R. K. Kumrular, "Automatic skin lesion segmentation," 2017 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), Istanbul, 2017, pp. 1-3
3. S. M. Jaisakthi, P. Mirunalini and C. Aravindan, "Automated skin lesion segmentation of dermoscopic images using GrabCut and k-means algorithms," in IET Computer Vision, vol. 12, no. 8, pp. 1088-1095, 12 2018.
4. P. Kharazmi, M. I. AlJasser, H. Lui, Z. J. Wang and T. K. Lee, "Automated Detection and Segmentation of Vascular Structures of Skin Lesions Seen in Dermoscopy, With an Application to Basal Cell Carcinoma Classification," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 6, pp. 1675-1684, Nov. 2017.
5. L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham and D. Feng, "Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks," in IEEE Transactions on Biomedical Engineering, vol. 64, no. 9, pp. 2065-2074, Sept. 2017.
6. Y. Yuan, M. Chao and Y. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," in IEEE Transactions on Medical Imaging, vol. 36, no. 9, pp. 1876-1886, Sept. 2017.
7. S. Sookpotharom, "Border Detection of Skin Lesion Images Based on Fuzzy C-Means Thresholding," 2009 Third International Conference on Genetic and Evolutionary Computing, Guilin, 2009, pp. 777-780.
8. Z. Ma and J. M. R. S. Tavares, "A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 2, pp. 615-623, March 2016.
9. J. Glaister, A. Wong and D. A. Clausi, "Segmentation of Skin Lesions From Digital Images Using Joint Statistical Texture Distinctiveness," in IEEE Transactions on Biomedical Engineering, vol. 61, no. 4, pp. 1220-1230, April 2014.
10. F. Peruch, F. Bogo, M. Bonazza, V. Cappelleri and E. Peserico, "Simpler, Faster, More Accurate Melanocytic Lesion Segmentation Through MEDS," in IEEE Transactions on Biomedical Engineering, vol. 61, no. 2, pp. 557-565, Feb. 2014.
11. B. S. Lin, K. Michael, S. Kalra and H. R. Tizhoosh, "Skin lesion segmentation: U-Nets versus clustering," 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, 2017, pp. 1-7.
12. E. Mendi, C. Yogurtcular, Y. Sezgin and C. Bayrak, "Automatic mobile segmentation of dermoscopy images using density based and fuzzy c-means clustering," 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Lisboa, 2014, pp. 1-6.