



Department of Computer Science and Engineering
University of Texas at Arlington
Arlington, TX 76019

Mining Skin Lesion Images with Spatial Data Mining Methods

Wenzhao Guo
wxg4124@omega.uta.edu

Y. Alp Aslandogan
alp@cse.uta.edu

Technical Report CSE-2003-19

Mining Skin Lesion Images with Spatial Data Mining Methods

Wenzhao Guo, Y. Alp Aslandogan

Department of Computer Science and Engineering

The University of Texas at Arlington

Contact Author: alp@cse.uta.edu

Abstract

We address the problem of identifying color variation in color skin lesion color. We adapt a spatial data mining method to this task and integrate with a segmentation method to identify significant color regions in an image. The resulting regions are compared with human perception via Kappa statistical test. Evaluation of the results indicate that the method approximates human judgment well and can be used as an automatic tool for mining skin lesion images. The approach is applicable to similar problems such as texture region identification and mining of other types of images.

1 Introduction

Melanoma is a lethal skin-cancer, but curable if detected early enough. Clinical features of melanoma are summarized as what's called ABCD: asymmetry, border irregularity, color variegation and diameter greater than 6mm. Early recognition of changes of lesion in terms of the above features provides important diagnostic and prognostic information.

Computer-aided analysis of the clinical features of skin lesions has been of significant importance and interest as a useful tool for early detection of the disease. This has been a very active research topic for decades. This paper will focus on the detection of color variation within the lesion since varied shades of brown, tan, or black are often the first signs of melanoma and as melanoma progresses the colors red, white and blue may appear.

From standard image processing perspective, image segmentation plays an important role in identifying lesions and sub-regions inside the lesions because accurate description and measurement of image features cannot be achieved without accurate image segmentation. A wide range of algorithms have been used for image segmentation, broadly categorized as pixel-based segmentation, region-based segmentation and edge detection.

This paper approaches the problem of image segmentation and color variation analysis of skin lesions from data mining perspective by applying a clustering technique. We have implemented a clustering technique, DBSCAN (Density Based Spatial Clustering of Applications with Noise), to segment skin lesion images, identify sub-regions inside lesions and extract color features.

2 Background

Due to its potential contribution to early detection of skin cancer, researches have been very active on image analysis of skin lesions in the past decades. Segmentation is one of the most popular standard image processing methods that are used for this purpose. The segmentation technique used in skin lesion analysis is built upon the existing generic image segmentation algorithms with some modifications that address the specific properties of skin lesion images.

Split-and-merge is a widely used region-based segmentation algorithm. It was first proposed by Horowitz and Pavlidis in 1976 [1]. Since then, modification and improvement have been made. Chen and Lin [2] proposed an adaptive split-and merge algorithm by learning domain-specific knowledge based on characteristic features and a hypothesis model. To address the problem that the edges of segments can only have horizontal and vertical orientations Wu's work [3] optimized the split-and-merge process by piecewise least-square approximation of an image intensity function. Manousakas et al. [4] made another improvement by applying simulated annealing

principles and controlled boundary elimination to traditional split-and-merge algorithm. Fuzzy technique is another method that has received extensive attention in pattern recognition and image segmentation. Lim and Lee proposed an algorithm based on the thresholding and fuzzy c-means techniques [5]. Chun and Yang presented a new and robust split-and-merge methodology using generic algorithm with a fuzzy measure [6].

Round et al's work on color segmentation for skin lesions is basically an application of split-and-merge algorithm [7]. Schmid and Fischer presented a color segmentation scheme to segment lesion images based on two-dimensional histogram analysis and fuzzy c-means clustering technique [8]. This scheme is actually an extension of Lim and Lee's method described in [5]. Umbaugh et al. transformed the RGB space to a spherical domain and developed the color image segmentation method based on a mathematically optimal transformation, the principal components transform (PCT) [9, 10]. A multi-channel segmentation algorithm was proposed by using both gray-level intensity and texture-based features [11, 12]. Hance et al. [13] compared six different color segmentation algorithms and found that PCT/median cut and adaptive thresholding algorithms provide the lowest average error and show the most promise. Recently, Xu et al. presented an automatic method for segmentation of skin lesions with focus on boundary detection [14]. Transformation of a color image into an intensity image and mapping of image intensities to enhance lesion boundary are reported as the main contribution of their work.

The reason why the segmentation method is popular is that it is able to partition the image into segments such that points or small regions inside one segment are homogeneous based on certain criteria and points or small regions from different segments are as different as possible, and as a result important properties of images can be identified. This is eventually in accordance with the basic concepts and principles of cluster analysis. Intuitively, data mining techniques, especially spatial clustering algorithms, are very likely to work fine to identify and extract features and attributes of images if we consider digital images as a data set of pixels or small regions that are homogeneous in terms of a certain color feature.

It has been reported that cluster analysis be used in such applications as pattern recognition, data analysis, image processing, and market research. However, based on our research, it is unknown to us that clustering techniques have been reported to mine skin lesion data. Given our problem domain, we are particularly interested in density-based methods, represented by DBSCAN [15].

DBSCAN is designed to discover clusters and noise in a spatial database [15]. It has two key parameters: Eps and MinPts. The neighborhood of an object within a radius Eps is called the Eps-neighborhood of the object. If the Eps-neighborhood of an object contains at least MinPts number of objects, then the object is called a core object. To find a cluster, DBSCAN starts with an arbitrary object o in the database. If the object o is a core object w.r.t. Eps and MinPts, a new cluster with o as the core object is created. DBSCAN continues to retrieve all density-reachable objects from the core object and add them to the cluster.

Two years later, GDBSCAN (Generalized DBSCAN) generalized the two key parameters of DBSCAN algorithm such that it can cluster point objects and spatially extended objects according to both spatial and non-spatial attributes [16]. The neighborhood of an object now is defined by a binary predicate NPred on a data set that is reflexive and symmetric. If NPred is true, then the neighborhood of an object is called NPred-neighborhood of an object. In other words, NPred-neighborhood of an object is a set of objects, S , which meet the condition that NPred is true. Corresponding to MinPts, another predicate, MinWeight of the set of objects, S , is defined such that it is true iff weighted cardinality for the set, $wCard(S)$, is greater or equal to minimum cardinality, MinCard, i.e. $wCard(S) \geq MinCard$. In search for clusters, GDBSCAN begins from an arbitrary object o . First, find out NPred-neighborhood of object o . Second, test MinWeight for the NPred-neighborhood of object o . If it is true, a new cluster with object o as a core object is created w.r.t. NPred and MinWeight. In such a way, GDBSCAN expands the concepts of the algorithm and broadens the scope of problem domains. This provides us with a basis on which we specialize the algorithm and apply it on lesion image segmentation.

Different from the previous image segmentation processing methods that pay much attention to improvement of accuracy of lesion boundary detection, the approach adopted in this paper focuses on color region recognition and mines color features of skin lesions by applying DBSCAN/GDBSCAN technique. It has three features: (1) It does not stop at recognition of lesions but goes inside lesions for recognition of sub-regions within lesions. The sub-regions obtained are represented as objects, which provides convenience for manipulation, retrieval and mining of color information. (2) Actions were taken to make the output (the number of sub-regions) produced by the computer program is as close to human perception as possible. (3) The color analysis processes implemented in this work are unsupervised. The program automatically chooses and adjusts related processing parameters based on the color information embedded in the lesion images.

3 Mining Skin Lesion Data with DBSCAN/GDBSCAN Technique

There are two major steps involved. First, the program splits the image into smaller regions until all the regions meet the homogeneity criteria set by the threshold for splitting. Second, the program groups the small split regions to form final regions of interest by DBSCAN algorithm. In other words, the image is segmented into disjoint regions that correspond to background skin and sub-regions inside lesion.

Euclidean distance in RGB color space is used to measure color distance and test homogeneity between color regions. There are no color transformations. The advantage associated with this is that once the objects of interest are identified they are represented in the original color space and are ready for retrieval and manipulation of color data.

The image is segmented in iterations. The aim of the first iteration is to identify the lesion. During the following iterations, the program moves inside the lesion and identifies sub-regions of the lesion. This approach allows the program to set different

parameters for different stages of segmentation. The program is designed in such a way that it automatically sets the parameters based on the color information of the whole image and/or the lesion and dynamically alters the parameters when it proceeds from one iteration to another.

Each segment identified by the program is an object that consists of a list of smaller objects (split regions) generated by the process of splitting. This provides us with an easy way to manipulate the color data of sub-regions inside lesion and produce quantitative output that characterizes the color variegation features.

Figure 1 shows the flow chart of the whole process of DBSCAN clustering.

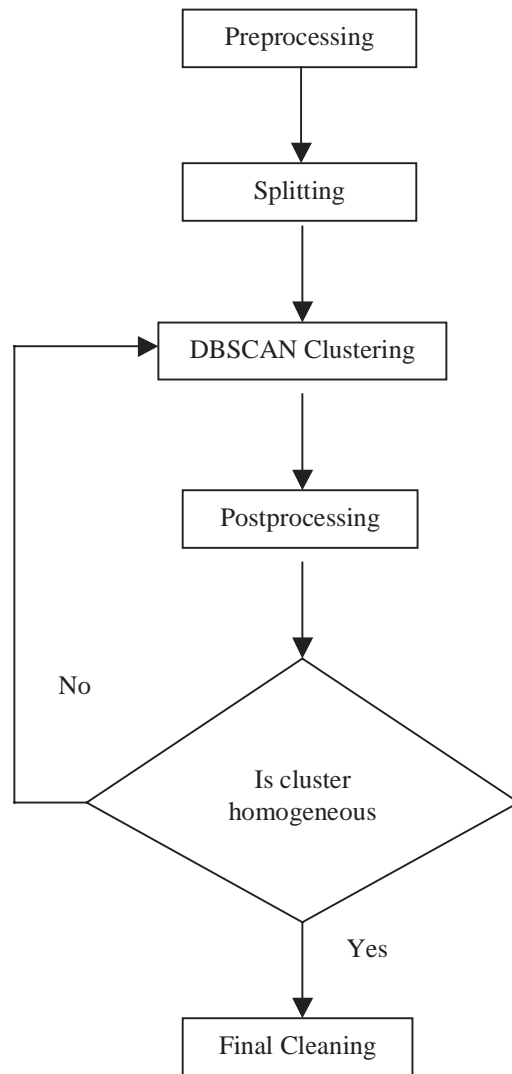


Figure 1. Flow chart of DBSCAN clustering

3.1 Preprocessing

Each image is smoothed before it goes to segmentation. The purpose for preprocessing is to remove noise such as reflection. Schmid and Fischer's work [8] reported that median filtering technique produces good smoothing result. We use the same method with a 5x5 neighborhood chosen. The advantage of median filtering is that it preserves the edges of color images while smoothing the images.

3.2 Process of Splitting

Splitting is a top-down process. A region is split into four sub-regions if it fails to pass homogeneity test. We start it from the highest-level region, which is the whole image we want to segment initially. The colors of a region and its four possible four sub-regions are represented by their mean color. The Euclidean distances between a region and each of its possible four child regions are computed. If any of the color distances obtained is greater than the threshold for splitting, the region will be split. Otherwise, the region remains unchanged. The process of splitting goes recursively until all the regions are found homogeneous or are too small to be split.

The threshold for splitting is a very important parameter. If it is set too high the accuracy of segmentation will decrease. The result will show rectangular borders around segments. On the other hand, if it is set too low, an image might be over-segmented and the size of split regions might be as small as just a few pixels. As a result, it might waste computation time and space since a smaller number of regions would suffice to produce equally good result. The program automatically determines the threshold for splitting as follows.

First, we get background color of the image. It is done by selecting four windows of 4x4 pixels from four corners of the image. We choose the median color of pixels as the background color. Such a treatment prevents interference of hair. Secondly, we divide the image into 100 sub-regions and use each of the sub-regions as a sample of the image. Third, we calculate the color distances between the background color and each of the samples and sort the color distances obtained. A big increase indicates a significant gap between two color regions, which the program needs to recognize and separate. The threshold for splitting is therefore chosen from those significant increases. Let's use Figure 2 to demonstrate how the threshold is obtained. There is a significant increase from sample n to sample $n+1$. If the increase is the first one that is greater than a predefined criteria and distance of sample n is also greater than a predefined criterion, we let the threshold for splitting to be the average of distances of

sample n and sample $n+1$. The two predefined criteria are set based on experimental testing on training image set and are independent of individual images. For an image that has less obvious color variation we need to choose a smaller threshold in order to recognize the color difference in the image while for an image that has obvious color variation a higher threshold will be determined and it would suffice to recognize the color difference.

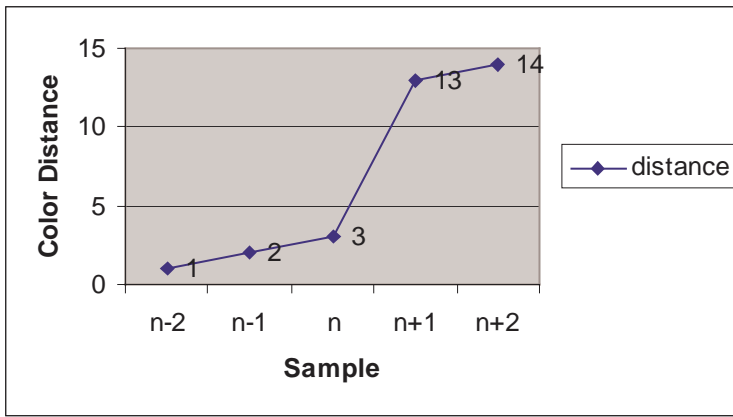


Figure 2. Demonstration on selection of threshold for splitting

Once the process of splitting is completed, a graph is constructed to represent the split regions and their connectivity. Two sub-regions are considered connected if they share a common boundary. The graph is implemented by an adjacency matrix.

3.3 DBSCAN Clustering to Form Segments

3.3.1 Selection of Two Key Parameters for DBSCAN Clustering

While GDBSCAN generalizes DBSCAN to a higher level to be able to deal with various problems in wider range, we will carry out specialization of GDBSCAN such that it fits into our domain. In our case, the data set that the algorithm is applied to is a set of split regions generated in the process of splitting and our objective is to cluster those split regions to form segments/objects of interest. Since each homogeneous region has a non-spatial attribute, color mean, and a spatial attribute, location in the

image, the first predicate NPred is defined as “color distance between two split regions is smaller than Eps AND the two split regions are connected” where Eps is a threshold of color distance that is obtained by the experimental testing on the training image set. For instance, given a split region reg_1 , another split region reg_2 is in the set of NPred-neighborhood of reg_1 iff the color distance between reg_1 and reg_2 is smaller than Eps AND reg_1 and reg_2 are spatially connected. Function $wCard(S)$ refers to total number of pixels of split regions involved. MinCard is a threshold of number of pixels, which is obtained experimentally. The predicate MinWeight is therefore defined as “the total number of pixels of split regions in the NPred-neighborhood, $wCard(S)$ is greater than MinCard.”

3.3.2 Process of DBSCAN Clustering

The program starts from an arbitrary split region reg_0 . First, based on the predicate NPred defined previously, find out all the connected regions that have a color distance of less than Eps. These selected regions form NPred-neighborhood of region reg_0 . Calculate $wCard(S)$, the total number of pixels of NPred-neighborhood. If the total number of pixels is greater than MinCard, a new cluster with region reg_0 as core object is created. Second, retrieve and add into the cluster all the possible directly density-reachable regions of each region selected in the first step. Third, the process goes on iteratively until no new region satisfies the predicate. This makes sure of the correctness of application of the algorithm, i.e. (1) for all region p, q , if region p belongs to a cluster and region q is density-reachable from region p w.r.t. NPred and MinWeight, then region q belongs to the same cluster; (2) for all region p, q in a cluster, region p is density-connected to region q w.r.t. NPred and MinWeight.

Let’s use Figure 3 as an example to demonstrate the process. Figure 3 shows a layout of split regions at the end of the process of splitting. Thick lines indicate the borders of split regions. Small cells indicate pixels. Clusters obtained are shown in different colors. Let MinCard be 10 pixels. Suppose we start from region₈. It has four connected regions: region₁, region₇, region₉ and region₁₃. But only region₁ and region₇

have a color distance of less than Eps from $region_8$. As a result, only $region_1$ and $region_7$ are in the NPred-neighborhood. Since the total number of pixels of $region_1$ and $region_7$ is greater than $MinCard$, 10, a new cluster with $region_8$ as core object is created. Next repeat the same procedure to find all possible directly density-reachable regions of $region_1$ and $region_7$ respectively. Starting from $region_1$, it has 5 adjacent regions: 2, 4, 6, 7, and 8. $Region_7$ and $region_8$ have been visited. $Region_2$ and $region_4$ have greater color distance than Eps . Only $region_6$ has smaller color distance than Eps and thus elected as NPred-neighborhood. Since $region_6$ has more than 10 pixels, it is directly density-reachable from $region_1$ and is added into the cluster. We cannot find any more directly density-reachable regions from either $region_1$, or $region_6$, or $region_7$, or $region_8$. As a result, a cluster consisting of $region_1$, $region_6$, $region_7$ and $region_8$ is formed.

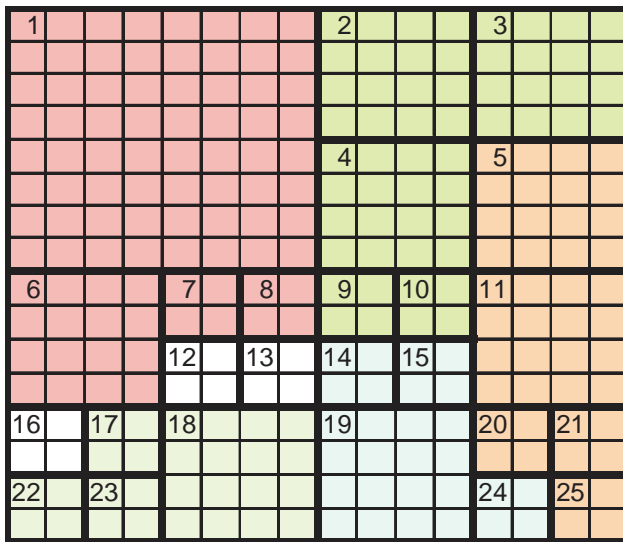


Figure 3. Demonstration of DBSCAN process

Figure 4 is a sample lesion image and Figure 5 is the result of initial DBSCAN process. The dark areas in Figure 5 show noise. They cannot be integrated into any cluster because they fail the predicate test. Notice that most of noise is along the borders of segments we want to identify. The explanation about this is that the thin area along

the borders of segments has much higher color variation. Thus, during the process of splitting, this area is divided into much smaller sub-areas, and during the process of clustering, smaller regions with higher color distance from their neighborhood are less likely to pass the predicate test to form or join a cluster.



Figure 4. Original image

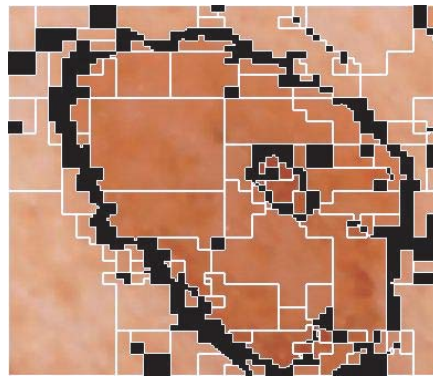


Figure 5. Image after the initial DBSCAN

3.3.3 Postprocessing

DBSCAN is an algorithm designed to discover clusters as well as noise in a spatial database. But in our case a non-noise label for each object of interest (segment) is required. Also as seen in Figure 5, the number of the clusters obtained from initial DBSCAN process is much more than the actual number of the clusters that we are interested in. Some of the clusters are just too small to be accepted as clusters. Therefore, postprocessing is necessary.

First, we merge the noise into the connected cluster that has smallest color distance from it. The noise is unclassified split regions in the implementation. It is likely that a noise region is surrounded by other noise regions and does not have a direct connection with any of clusters. These noise regions may not be able to be merged until some of their surrounding noise regions get merged first. Thus, it may take a few iterations to have all noise regions merged.

Second, we merge clusters with the connected cluster that have smallest color distance from it by a process of refinement. Here we need to set a threshold to control the process. It can be either a minimal number of clusters that are produced as final output or a threshold of color distance. Obviously, a threshold of color distance is more appropriate. If the smallest color distance between a cluster and its connected neighbors is greater than the threshold it won't be merged with any cluster any more.

The threshold for refinement is more relaxed compared with those used in the processes of splitting and initial clustering and is set by multiplying the threshold for splitting by a factor. The factor is dependent upon the significance of color variation among the initial segments. The latter can be represented by color mean absolute deviation of the initial segments. Normally the area with less significant color variation normally has smaller color mean absolute deviation. Thus the threshold for refinement obtained here is smaller and the criterion for merging becomes stricter. With stricter threshold the program is able to detect less significant color variation. The threshold used in process of refinement is automatically determined by the program.

Figure 6 shows the result up to this point.

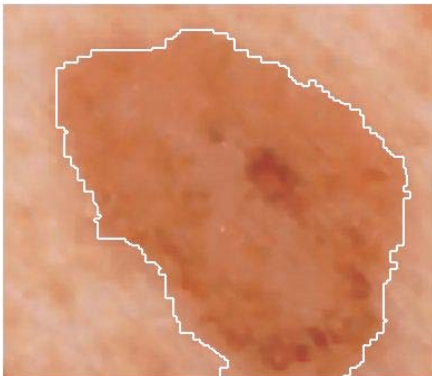


Figure 6. Image after the first iteration

3.3.4 More Iterations of DBSCAN Clustering

Now we have finished the first iteration of DBSCAN clustering and achieved the first objective that is to identify the lesion from the background skin. Our next goal is to get inside the lesion to identify sub-regions if they exist.

We set up three conditions to test whether or not a cluster produced by the first iteration of clustering needs more iterations of DBSCAN clustering: (1) The size of the cluster needing second DBSCAN clustering is relatively large. Normally it should be larger than one tenth of the original image in terms of the number of pixels. (2) The color distance between the cluster that needs second DBSCAN clustering and the background color should be large than a threshold. (3) The cluster that needs second DBSCAN clustering has relatively significant color variation. This is represented by the color mean absolute deviation of split regions within that cluster. If its color mean absolute deviation is greater than a threshold, the cluster is seen to have significant color variation. If a cluster meets any of the three conditions listed above, another iteration of DBSCAN clustering will apply.

During the second DBSCAN clustering, both Eps and MinCard increase by a small portion because segment (cluster) generated by the first DBSCAN clustering normally has less significant color variation than the original image. Stricter criteria help to recognize less significant difference.

In order to avoid over-segmentation, additional DBSCAN clustering won't exceed two iterations.

3.3.5 Final Cleaning Process

There are two actions that may be taken at this step.

One of our main goals is that the result produced the computer program is as close to human perception as possible. Normally it is quite difficult for human eyes to recognize small color differences. Also human perception tends to favor color regions with relatively visible size and border. As a result, the number of colors/sub-regions inside lesion tends to be relatively small. On the other hand, the computer program has more powerful ability to detect tiny color differences not detected by human eyes. Therefore, the computer program tends to find larger numbers of colors/sub-regions inside lesions especially in the thin band along the border of lesion. It may be necessary to take an extra merging process with more relaxing threshold to close the gap between

the computer program and human perception if too many clusters are shown in the final output.

Another scenario is that sometimes there exist a few tiny segments at the end of previous processes. We regard segments as noise if they are smaller than two hundredth of the size of the original image in terms of number of pixels. Those tiny segments will be forced to merge with their nearest neighborhood cluster.

Figure 7 shows the output at the end of the whole DBSCAN process.



Figure 7. Image after the second iteration

3.4 Experimental Results

3.4.1 Lesion Image Segmentation

In order to test the correctness and efficiency of the program and set up values for related parameters, the program was first applied to 18 images, which were chosen from a 500-image database. The selection of the 18 images was done in the way that they could be as representative as possible. It includes images with various sizes, with various colors, high color variation and low color variation, with lesions having clear border and blur border, dark and light, with hairs and without hairs, with reflection and other types of noise, etc.

As explained earlier, the threshold for splitting is chosen by the program based on the color information of images. Therefore, the thresholds for splitting are different from images to images. The actual threshold is a square one. By doing so, we may not need to calculate square root when we calculate Euclidean distance between two color regions. Main parameters are set up as follows.

Eps: $\text{thresholdS} / 3$, where thresholdS is the threshold for splitting

MinCard: 80 pixels

Experiments show that the program may not work very well if the image is very hairy, or there is a relatively large area of reflection.

Next, the program was applied to 117 images. The result was shown with more accurate boundaries and fewer rectangular edges compared with traditional region merging methods.

3.4.2 Correlation and Agreement between DBSCAN Clustering and Human Perception

Experiments were carried out to test the correlation and agreement between the computer program and human perception. 5 observers participated in the experiment. Each observer independently reviewed the images and no observer had seen the results produced by the computer program before she/he reviewed the images.

We choose the number of colors/sub-regions inside lesion as the attribute for the evaluation of computer program and use majority vote of the observations for human perception. We convert the evaluation originally represented by the number of colors/sub-regions inside lesion into numeric ranks, shown in Table 1. In our experiments, since no human observation identifies an image with more than 3 colors/sub-regions inside lesion and the computer program only identifies 6 out of 117 images, which have 4 colors/sub-regions inside lesion (no case has more than 4 colors/sub-regions), we rank the result that has equal to or greater than 3 colors/sub-regions as 3.

Table 1. Conversion of original evaluation into numeric ranks

Number of Colors/Sub-regions	Rank
1	1
2	2
3+	3

(a) Correlation between the DBSCAN Clustering and human perception

Correlation coefficient is 0.30, indicating a fair but positive correlation. Hypothesis test (t tests) [17] for the correlation coefficient shows significant positive correlation with 95% confidence level.

(b) Agreement between the computer program and human perception

Kappa coefficient of Cohen [18] is recommended as an indicator of observer agreement. If Kappa coefficient is equal to zero, it indicates no agreement. A negative Kappa coefficient means disagreement whereas a positive one means agreement [19].

We apply Kappa coefficient test on the agreement between the computer program and human perception.

Table 2 summarizes the results from the computer program and human perception.

Table 2. Summary of evaluation by DBSCAN and human perception

		Human Observation			Total
		1	2	3	
Computer Program	1	2	3	1	6
	2	3	58	5	66
	3	2	28	15	45
Total		7	89	21	117

The result of agreement test: Kappa coefficient is 0.2803, with 95% confidence interval of [0.1318, 0.4288], showing a significant agreement between the computer program and human perception.

4 Conclusion

Image segmentation and color feature analysis of skin lesions is not a new topic. But this paper adopts a new approach to this problem by combining a cluster analysis algorithm, DBSCAN, with image processing techniques. The main contribution of this work is the successful application of the clustering algorithm in skin lesion image analysis.

The result of segmentation by DBSCAN clustering was shown with more accurate boundary and fewer rectangular edges compared with traditional split-and-merge method. Experiments have also indicated that there are positive and statistically significant correlation and agreement between the outputs of the computer program and human perception.

Further work needs to be done to address the influence of noise, such as hairs and reflection, so as to increase the robustness of the program.

References

- [1] Horowitz, S.L., and Pavlidis, T. Picture Segmentation by a Tree Traversal Algorithm, *JACM*, Vol. 23, No. 2, April 1976, page(s) 368-388.
- [2] Chen, S.Y, and Lin, W.C. Split-and-merge image segmentation based on localized feature analysis and statistical tests, *Graphical Models and Image Processing*, Vol. 53, No. 5, Sep 1991, page(s) 457-475.
- [3] Wu, X. Adaptive split-and-merge segmentation based on piecewise least-square approximation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No. 8, Aug 1993.
- [4] Manousakas, I.N., Undrill, P.E., Cameron, G.G., and Redpath, T.W. Split-and-merge segmentation of magnetic resonance medical images: performance evaluation

and extension to three dimensions, *Computers and Biomedical Research*, Vol. 31, 1998, page(s) 393-412.

[5] Lim, Y.W., and Lee, S.U. On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques, *Pattern Recognition*, Vol. 23, No. 9, 1990, page(s) 935-952.

[6] Chun, D.N., and Yang, H.S. Robust image segmentation using genetic algorithm with a fuzzy measure, *Pattern Recognition*, Vol. 29, No. 7, 1996, page(s) 1195-1211.

[7] Round, A.J., Duller, A.W.G., and Fish, P.J. Color segmentation for lesion classification, *Proceedings – 19th International Conference – IEEE/EMBS Oct 30 – Nov 2, 1997 Chicago, IL. USA*

[8] Schmid, P., and Fischer, S. Colour segmentation for the analysis of pigmented skin lesions.

[9] Umbaugh, S.E., Moss, R.H., and Stoecker, W.V. "Automatic color segmentation of images with application to detection of variegated coloring in skin tumors" *IEEE Engineering in Medicine and Biology Magazine*, Vol. 8, No. 4, Dec. 1989, Page(s): 43 –50.

[10] Umbaugh, S.E., Moss, R.H., Stoecker, W.V. and Hance, G.A. "Automatic color segmentation algorithms with application to skin tumor feature identification" *IEEE Engineering in Medicine and Biology Magazine*, Vol. 12, No. 3, Sept. 1993, Page(s): 75 –82.

[11] Hance, G.A., Umbaugh, S.E., Moss, R.H., and Stoecker, W.V. "Unsupervised color image segmentation: with application to skin tumor borders" *IEEE Engineering in Medicine and Biology Magazine*, Vol. 15, No. 1, Jan.-Feb. 1996, Page(s): 104 –111.

[12] Dhawan, A. P., and Parikh, M. "Knowledge-based color and texture analysis of skin images" *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 12, No. 3, 1990.

[13] Dhawan, A.P., and Sicsu, A. Segmentation of images of skin lesions using color and texture information of surface pigmentation, *Computerized Medical Imaging and Graphics*, Vol. 16, No. 3, 1992, page(s) 163-177.

- [14] Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., Dhawan, A., and Huntley, A. Segmentation of skin cancer images, *Image and Vision Computing*, Vol. 17, 1999, page(s) 65-74.
- [15] Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. A density-based algorithm for discovering cluster in large spatial databases, *Proc 1996 Int. Conf. Knowledge Discovery and Data Mining (KDD'96)*, Portland, OR, 1996, page(s) 226-231.
- [16] Sander, J., Ester, M., Kriegel, H.-P., and Xu, X. Density-based clustering in spatial databases: the algorithm DGBSCAN and its applications, *Data Mining and Knowledge Discovery, an Int. Journal*, Kluwer Academic Publishers, Vol. 2, No. 2, 1998, page(s) 169-194.
- [17] Rees, D.G. *Essential Statistics*, 4th ed., Chapman & Hall/CRC, 2001.
- [18] Cohen, J. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, Vol. 20, No. 1, 1960.
- [19] Woolson, R.F. *Statistical Methods for the Analysis of Biomedical Data*, John Wiley & Sons, 1987.