Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007)

Parameterization of Dermoscopic Findings for the Internet-based Melanoma Screening System

Hitoshi Iyatomi∗†, Hiroshi Oka‡, M.Emre Celebi ⋄, Masaru Tanaka⋆, and Koichi Ogawa∗
∗ Department of Electrical Informatics, Hosei University Faculty of Engineering, Tokyo, JAPAN.
† Department of Dermatology, Keio University School of Medicine, Tokyo, JAPAN.
‡ Department of Information and Computer Science, Keio University Faculty of Science and Technology, Yokohama, JAPAN.
⋄ Department of Computer Science and Engineering, University of Bridgeport, CT, USA.
⋆ Department of Dermatology, Tokyo Women's Medical University Medical Center East, Tokyo, JAPAN.
∗ 3-7-2 Kajino-cho Koganei, 184-8584, Tokyo, JAPAN.
∗ iyatomi@k.hosei.ac.jp

***Abstract*— In this paper, we develop an automated method to recognize the dermoscopic criteria used for diagnosing melanomas as defined by two commonly used diagnostic schemes, namely the ABCD rule and the 7-point checklist. We use a database of 105 dermoscopy images and their dermoscopic findings determined by four dermatologists as the gold standard. We extract 356 objective image parameters from the images and build multiple regression models for the 15 clinical findings defined by the aforementioned diagnostic schemes. Our model provides comparable results to dermatologists in recognizing almost all of the findings. The results of our preliminary experiments show that computer-based image analysis has the potential of identifying the dermoscopic criteria.**

***Index Terms*— dermoscopy, computer-aided diagnosis (CAD), melanoma, ABCD rule, 7-point checklist, feature extraction**

I. Introduction

In the past few decades, the incidence of malignant melanoma has increased gradually in most parts of the world. In Australia, the incidence of melanoma is now approaching 50 cases per 100,000 population [1]. Although advanced malignant melanoma is often incurable, early-stage melanoma can be cured in many cases, particularly before the metastasis stage. For example, patients with a melanoma less than or equal to 0.75 mm in thickness have a good prognosis and their five-year survival rate is greater than 93%[2]-[5]. Therefore, early detection is essential for the reduction of melanoma-related deaths.

In 1987, Soyer *et al.*[6] introduced Dermoscopy, a non-invasive skin imaging technique that uses optical magnification and either liquid immersion and low angle-of-incidence lighting or cross-polarized lighting to make the contact area translucent, making subsurface structures more easily visible when compared to conventional macroscopic (clinical) images. Later, Stolz *et al.* developed the ABCD rule [7] and Argenziano *et al.* developed the 7-point checklist [8] as convenient diagnostic schemes to be used with dermoscopy images in which morphological characteristics of a pigmented skin lesion (PSL) such as asymmetry, border sharpness, and color variegation are quantified and a diagnosis is made based on the total score. Mayer reported that dermoscopy improved the sensitivity by 10-27% between 1983 and 1997 [9]. However, dermoscopic diagnosis is often subjective and is therefore associated with poor reproducibility and low accuracy especially in the hands of inexperienced dermatologists. Despite the use of dermoscopy, the accuracy of expert dermatologists in diagnosing melanoma is still estimated to be about 75-84% [1][10][11].

Several groups have developed automated analysis procedures to overcome these problems and reported high levels of diagnostic accuracy [12]-[16]. Blum *et al.* [13] reported a sensitivity (SE: melanoma detection accuracy) of 82.2% and a specificity (SP: benign detection accuracy) of 82.5% on 837 cases of suspicious melanocytic lesions using artificial neural networks. Rubegni *et al.* [14] achieved a SE of 94.3 % and a SP of 93.8% on 350 cases of nevi and 200 cases of malignant melanoma. However, a significant problem with these approaches has persisted. These studies were designed to develop a screening system for new patients using standalone systems and therefore they have not been open to the public.

We have developed an Internet-based melanoma screening system [17]. The URL of the website is http://dermoscopy.soft.ics.keio.ac.jp. Fig. 1 shows the schematic of our screening system. When one uploads a dermoscopy image and the associated clinical data to the system, the system extracts the tumor area, calculates the tumor characteristics and reports a diagnosis based on linear discriminant analysis. As of April 2006, the latest version of our system is equipped with the dermatologist-like tumor-area extraction algorithm that achieved superior extraction performance [18] and a neural network classifier. The system achieved 90.7% SE and 80.7% SP with a leave-one-out cross-validation strategy on a set of 319 dermoscopy images. Our present system provides the final diagnosis results in the form of a malignancy score between 0 and 100. It is desirable that a system provides the grounds for the diagnostic results in accordance with common clinical findings, such as those defined by the ABCD rule or the 7-point checklist. However, since these dermoscopic findings are defined subjectively, their automated quantification is difficult. Although several researchers attempted to extract these features with image

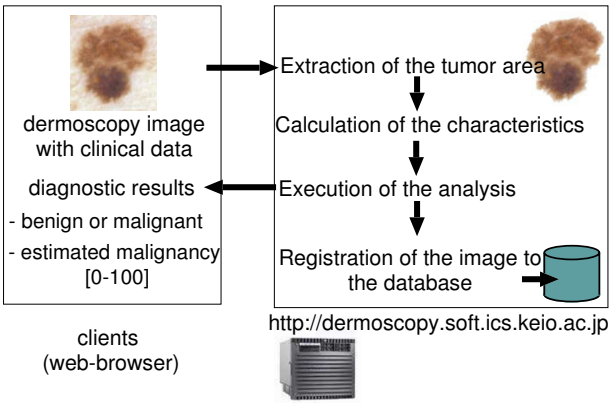

Fig. 1. Schematic of the Internet-based screening system.

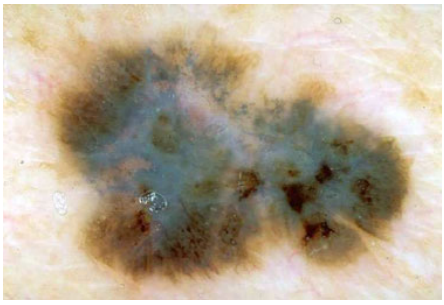

Fig. 2. Sample of dermoscopy image. (melanoma)

processing techniques [19]-[21], no general solution has been proposed especially for the extraction of structural features such as pigment networks, streaks etc.

In this paper, we develop an automated method that recognizes commonly used dermoscopic criteria defined by the ABCD rule and the 7-point checklist. Our method parameterizes the characteristics of PSLs objectively and calculates the probability of the corresponding finding from these parameters automatically. Therefore, it does not require the precise extraction of the image structures that are subjectively defined by the dermoscopic criteria.

II. Materials and Methods

The objective of this research is to build a computer-based system that recognizes dermoscopic findings as accurately as expert dermatologists. In this study, we focused on the ABCD rule [7] and the 7-point checklist [8] and built multiple regression models for a total of 15 findings defined by these diagnostic schemes.

A. Materials

Digital dermoscopy images of PSLs were collected from two university hospitals (University of Naples, Italy, and University of Graz, Austria). These were 24-bit images with a typical resolution of 768x512 pixels. In order to achieve scale-invariant results, the images were reduced to 40 pixels/mm. All of the cases were diagnosed based on histopathological examination of the biopsy material. Since we had no control over the image acquisition and camera calibration, images that satisfied at least one of the following criteria were omitted from the study: (i) the lesion does not fit entirely within the image frame, (ii) the lesion is part of an acral or mucosal area, and (iii) presence of too much hair. This selectivity was necessary in order to ensure accurate border detection and reliable feature extraction.

A total of 105 images free from the abovementioned problems were included in this study. The diagnosis distribution was as follows: 65 benign nevi (40 Clark nevi, 15 Reed nevi, and 10 blue nevi) and 40 melanomas (including 16 cases of melanoma in situ). For each image, four experienced dermatologists determined the dermoscopic findings according to the ABCD rule and the 7-point checklist.

B. Clinical items for diagnosing dermoscopy image

We briefly introduce the ABCD rule and the 7-point checklist here.

1) ABCD rule: This is one of the most well-known semi-quantitative diagnosis scheme. It quantifies the asymmetry (A), border sharpness (B), color variegation (C) and the number of differential structures (D) present in the lesion. Table I summarizes these findings and their relative weights. A represents the degree of asymmetry of the tumor. Assuming a pair of orthogonal symmetry axes intersecting at the centroid of the tumor, A can be 0 (symmetry on both axes), 1 (symmetry on one axis), or 2 (no symmetry). B represents the number of border octants with a sharp transition. C represents the number significant colors present in the tumor. Six colors are considered significant: white, red, light-brown, dark-brown, blue-gray, and black. D represents the number of differential structures (pigment network, structureless or homogeneous areas, streaks, dots, and globules) present in the tumor. Using the ABCD rule, the total dermoscopy score (TDS) is calculated as follows:

$$TDS = (A \times 1.3) + (B \times 0.1) + (C \times 0.5) + (D \times 0.5). \quad (1)$$

TDS below 4.75 indicates benignity, whereas TDS above 5.45 indicates malignancy. A score between these limits corresponds to a suspicious case that requires clinical follow-up.

As an example we shall calculate the TDS for the dermoscopy image shown in Fig.2. Since neither the long nor the short axis show symmetry property, A = 2. It can be seen that 5 of the 8 border octants exhibit a sharp transition, so B = 5. Five significant colors and four differential structures are observed, so C = 5 and D = 4. When we substitute these values into the above equation, we obtain a TDS value of 7.6 (> 5.45) which indicates that this lesion is malignant.

2) 7-point checklist: This is another well-known diagnostic method that requires the identification of only 7 dermoscopic structures that are shown in Table II. The score for a lesion is determined as the total weight of the structures present in it. If this score is greater than or equal to 3, then the lesion is considered to be malignant. For the case of Fig.2, blue-whitish veil (2 point), irregular streaks (1 point), irregular pigmentation (1 point), irregular dot (1 point) and regression (1 point) are

TABLE I
ABCD RULE CRITERIA

Criterion	Description	score	weight
Asymmetry	Number of asymmetry axes	0 - 2	× 1.3
Border	Number of sharp border octants	0 - 8	× 0.1
Color	Number of significant colors	1 - 6	× 0.5
Differential structure	Number of specific structures	1 - 5	× 0.5

TABLE II
7-POINT CHECKLIST CRITERIA

Major criteria	weight
1. Atypical network	2
2. Blue-whitish veil	2
3. Atypical vascular pattern	2
Minor criteria	
4. Irregular streaks	1
5. Irregular pigmentation	1
6. Irregular dots/globules	1
7. Regression structures	1

observed and therefore (total 6 points) this lesion is considered to be malignant.

3) Building the regression model: For each image, four dermatologists determined the dermoscopic findings defined by the ABCD rule and the 7-point checklist. Whenever a dermatologist recognized a specific dermoscopic structure in an image, a teach signal of 1 is assigned to the image and otherwise a 0 is assigned. Note that for "asymmetry", "border" and "color", the range of these findings are 0 - 2, 0 - 8 and 1 - 6, respectively and therefore the number was directly used as the value of the teach signal. We calculated the weighted average of these teach signals considering the dermoscopy experience of the dermatologist and used them as objective variables of regression.

For the 105 dermoscopy images, we extracted the tumor area using the dermatologist-like tumor extraction algorithm [18] and calculated a total of 356 parameters. These parameters can be roughly categorized as color, symmetry, border and texture related properties. As color related parameters, we calculated minimum, average, maximum, standard deviation, skewness and the number of colors in the RGB and HSV color spaces quantized to 8^3 and 16^3 colors, respectively. These 34 parameters were also calculated from the peripheral part of the tumor that is defined as the outer region of the tumor with an area equal to 30% of the tumor area. Consequently, a total of 68 color related parameters were calculated.

As the symmetry parameters, a total of 80 parameters were calculated. We designed 10 thresholds of intensity starting from 20 up to 225 with a step size of 25. Within the tumor area, thresholding was performed and the areas with an intensity lower than the threshold were determined. For each thresholded region, we calculated eight parameters: the ratio between the area of the region and the tumor area, circularity, distance between the center of gravity of the region and that of the tumor, standard deviation and skewness of the (x,y) distribution.

As the border parameters, a total of 32 parameters were calculated. The tumor area was divided to eight equi-angle regions and the mean value and the gradient of blue and luminance were calculated in each region.

In order to characterize the structure of a dermoscopy image, we calculated various texture parameters. We prepared 11 different co-occurrence matrices with distance values δ ranging from 1/2 to 1/64 of the length of the long axis of the tumor. Based on each co-occurrence matrix, energy, moment, entropy and correlation were calculated in four directions (0, 45, 90, and 135 degrees). Accordingly, a total of 11 × 16 = 176 texture parameters were calculated.

After the calculation of 356 parameters from each image, we calculated the correlation between these parameters and eliminated the highly correlated ones in order to ensure a robust regression model. Parameters, whose absolute correlation coefficient exceeds 0.99, were eliminated except for one representative. Next, we performed incremental stepwise feature selection with Wilks' lambda hypothesis test [22] to build multiple regression models for each clinical finding.

4) Evaluation criteria: We evaluated the accuracy of the built regression model with mean-absolute-error (MAE), determination coefficient adjusted by degree of freedom (R^2) and standard deviation of the diagnostic results of four dermatologists (σ_{doc}). We also conducted a t-test with the null hypothesis that the dermatologists and the regression model have the same recognition capability for the dermoscopic findings.

III. RESULTS

After the elimination of the highly correlated parameters, 163 parameters remained. The incremental stepwise feature selection method selected multiple regression models having 3 to 16 parameters. Table III summarizes the evaluation results of the built regression model. From left to right the columns correspond to the number of constituent parameters of the regression model ($\#par$), mean absolute error of the regression model (MAE), standard deviation of the diagnostic results of the four dermatologists (σ_{doc}), determination coefficient adjusted by the degree of freedom (R^2) and p value of the statistical t-test.

Considering that the ranges for "asymmetry", "border" and "color" are 0-2, 0-8 and 1-6, respectively, we see that our regression models achieved MAE less than 20-30% of the corresponding range.

IV. DISCUSSION

A. Regression Model

From Table III, it can be seen that our regression models achieved MAE about 20-30% of the range for the corresponding findings. Although 20-30% MAE is not quite small, this is less than or equal to the standard deviation of the diagnostic results of the dermatologists.

Table IV compares the recognition results for the dermoscopy image shown in Fig.2 by a dermoscopy training website on the Internet (Ref.[11]), expert dermatologist (expert),

TABLE IV

COMPARISON BETWEEN BUILT REGRESSION MODEL AND DERMATOLOGISTS - FOR FIG.2

	ABCD rule								7-point checklist							score	
	Asym	Bord	Col	Dpig	Dbra	Dstl	Ddot	Dglo	a-pn	bwv	a-vp	i-st	i-pg	i-dg	reg	ABCD	7point
Ref.[11]	2	5	5	1	0	1	1	1	-	-	-	-	-	-	-	7.6	N/A
expert	2	8	4	1	1	1	1	1	0	1	0	1	1	1	1	7.9	6
average	2.0	5.8	4.2	0.9	0.8	0.8	0.7	0.7	0.4	1.0	0.0	0.7	1.0	0.7	0.8	7.23	6.00
our model	2.17	5.47	3.91	0.80	0.61	0.84	0.90	0.70	0.47	1.13	0.05	0.74	0.91	0.81	0.64	7.25	6.40

TABLE III

COMPARISON BETWEEN BUILT REGRESSION MODEL AND DERMATOLOGISTS -SUMMARY

	criterion	#par	MAE	σ_{doc}	R^2	p
ABCD rule	A: asymmetry (0-2)	11	0.526	0.399	0.478	0.120
	B: border (0-8)	4	1.529	2.318	0.570	0.330
	C: color (1-6)	16	0.406	0.565	0.696	0.275
	D: pigment network	6	0.279	0.184	0.543	< 0.05
	D: branched streaks	8	0.240	0.196	0.572	0.327
	D: structure less	9	0.267	0.343	0.435	< 0.01
	D: dots	10	0.279	0.306	0.447	< 0.01
	D: globules	12	0.237	0.331	0.541	< 0.01
7-point checklist	atypical pigment net.	5	0.303	0.270	0.382	0.164
	blue-whitish veil	7	0.233	0.236	0.551	0.077
	atypical vascular str.	3	0.123	0.092	0.149	0.169
	irreg. streaks	6	0.213	0.170	0.496	0.309
	irreg. pigmentation	3	0.301	0.327	0.231	0.260
	irreg. dots	9	0.206	0.276	0.441	< 0.01
	regression	13	0.202	0.132	0.506	0.354

average of the four dermatologists (average) and our regression model (model). Each column represents the dermoscopic findings defined by the ABCD rule and the 7-point checklist. Although this is just a single example, we can observe a high variation among the dermatologists and, in spite of this our regression models provides reasonable results. Similar trends were observed in other cases. MAE for atypical vascular pattern and irregular pigmentation of the 7-point checklist could be kept smaller, but R^2 was also lower when compared to other models. In the former, only a small number of dermoscopy images with this pattern was included in image set so that most of objective variables (weighted average diagnosis by dermatologists) for this model were almost 0 except for a few cases. Because regression model was built to minimize the mean square error, it can be considered that the model was trained to predict almost 0 for all cases. In the latter, MAE of irregular pigmentation is smaller than σ_{doc}, but in this case itself has a large value. Because our model was built based on the average of the diagnostic results by dermatologists and they have a high variation, our model might not capture the essential characteristics of the dermoscopic findings adequately. In addition, the number of selected parameters by the stepwise method for these two regression models was only three. This indicates that our 163 basic parameters with a linear classifier could not simulate the recognition results of dermatologists adequately. We have to collect more dermoscopy images with the corresponding dermoscopic findings and find effective parameters to characterize them.

According to the results of the t-test, 10 out of 15 built regression models could not reject the null hypothesis at the risk factor of 0.05. This indicates that the recognition capability for these 10 dermoscopic findings by the dermatologists and the regression model can be considered to be statistically equivalent.

Note that we cannot immediately conclude that the built regression models can be used in practice because, in our experiments, the modeling and evaluation data were the same. However, we believe that our approach has the potential of identifying dermoscopic findings.

B. Effective parameters to express dermoscopic findings

One of the color related parameters, the number of colors in the tumor represented by the $16^3 = 4096$ RGB color model, was selected by seven regression models. They were "asymmetry", "color", "branched streaks", "dots", "globules" of the ABCD rule and "irregular streaks" and "irregular dot/globules" of the 7-point checklist. Furthermore, this parameter was selected in the top three by six of the seven models. The incremental stepwise method used in this experiment searches the parameter space with a round-robin strategy and finds the parameter that achieves the lowest MAE at each step. Therefore, the parameters selected earlier are more important than those selected later. Consequently, we confirmed again that the number of colors is important for recognizing the dermoscopic findings. In addition, this parameter represents the polychroism of melanomas. The parameter "the dark area distribution whose intensity is less than 30 along the long axis" is selected by five dermoscopic findings: "color", "pigment network", "dots", "blue-whitish veil" and "irregular streaks". It can be concluded that this parameter represents the blackness and non-uniformity of melanoma.

From these experiments, we found several important characteristics that represent the subjective dermoscopic findings. As a consequence of this research, our Internet-system will be able to provide not only diagnostic results, but also the grounds for the diagnosis - the quantitative score of common dermoscopic findings. We will further investigate this research and mount these regression models on our Internet-based system.

V. CONCLUSIONS

In this paper, we developed a recognition model for the dermoscopic findings defined by the ABCD rule and the 7-point checklist. We extracted a total of 356 objective parameters from a set of 105 dermoscopy images and built

multiple regression models using the ground-truth determined by four dermatologists. We demonstrated that our model is as accurate as expert dermatologists in recognizing almost all of the clinical findings. We also found that the number of colors and the distribution of darker areas are important features in the recognition of these findings. The results of our preliminary experiments show that computer-based image analysis has the potential of identifying dermoscopic criteria. In the near future, we will mount these models on our Internet-based melanoma screening system.

ACKNOWLEDGMENT

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), 17790788, 2005-2006.

REFERENCES

[1] W.Stolz, O.B.Falco, P.Bliek, M.Kandthaler, W.H.C.Burgdorf and A.B.Cognetta, "Color Atlas of Dermatoscopy – 2nd enlarged and completely revised edition," Berlin, Blackwell publishing, 2002.

[2] F.L.Meyskens Jr, D.H.Berdeaux, B. Parks, T.Tong, L.Loescher and T.E.Moon, "Natural history and prognostic factors incluencing survival in patients with stage I disease," Cancer, Vol.62, pp.1207-1214, 1988.

[3] P.Buttner, C.Garbe, J.Bertz, G.Birg, B. d'Hoedt, H.Drepper et.al, "Optimized cutoff points of tumor thickness and importance of Clark's level for prognostic classification," Cancer, Primary cutaneous melanoma, Vol.75, pp.2499-2506, 1995.

[4] C. Garbe, P.Buttner, J.Bertz, G. Berg, B.d'Hoedt, H.Drepper et.al, "Identification of prognostic groups and estimation of individual prognosis for 5093 patients," Cancer, Primary cutaneous melanoma, Vol.75, pp.2484-2491, 1995.

[5] C.M.Balch, T.M.Murad, S.J.Soong, A.L.Ingalls, P.C.Richards and W.A.Maddox, "Tumor thickness as aguide to surgical management of clinical stage I melanoma patients," Cancer, Vol.43, pp.883-888, 1979.

[6] H.P.Soyer, J.Smolle, H.Kerl and H.Stettnre, "Early diagnosis of malignant melanoma by surface microscopy," Lancet, Vol.2, p.803, 1987.

[7] W.Stolz, A.Riemann, A.B.Cognetta, L.Pillet *et al,* "ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma," European Journal of Dermatology, No.4, pp.521-527, 1994.

[8] G.Argenziano, G.Fabbrocini, P.Carli *et al.* "Epiluminescence microscopy for the diagnosis of ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis, " Archives of Dermatology, No.134, pp.1536-1570, 1998.

[9] J.Mayer, "Systematic review of the diagnostic accuracy of dermoscopy in detecting malignant melanoma," Med. Journal of Australia, Vol.167, pp.206-210, 1997.

[10] G.Argenziano H.P.Soyer, S.Chimenti et.al, "Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet," Journal of American Academy of Dermatology, Vol.48, No.5, pp.679-693, May, 2003.

[11] http://www.dermoscopy.org

[12] F.Ercal, A.Chawla, W.V.Stoecker, H-C.Lee and R.H.Moss, "Neural network diagnosis of malignant melanoma from color images," IEEE Trans. on Biomedical Engineering, Vol.41, No.9, pp.837-845, Sept. 1994.

[13] A.Blum, H.Luedtke, U.Ellwanger, R.Schwabe, G.Rassner and C.Garbe, "Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions," British Journal of Dermatology, Vol.151, pp.1029-1038, 2004.

[14] P.Rubegni, G.Cevenini, M.Burroni, R.Perotti, G.Dell'Eva, P.Sbano et.al, "Automated diagnosis of pigmented skin lesions," International Journal of Cancer, Vol.101, pp.576-580, 2002.

[15] K.Hoffman, T.Gambichler, A.Rick et al, "Diagnostic and neural analysis of skin cancer (DANAOS). A multicentre study for collection and computer-aided analysis of data from pigmented skin lesions using digital dermoscopy," British Journal of Dermatology, Vol.149, pp.801-809, 2003.

[16] M.Burroni, P.Sbano, G.Cevenini, M.Risulo, G.Dell'Eva, P.Barbini, C.Miracco, C.Miracco, M.Fimiani, L.Andreassi and P.Rugegni, "Dysplastic naevus vs. in situ melanoma: digital dermoscopy analysis," British Journal of Dermatology, Vol.152, pp.679-684, 2005.

[17] H.Oka, M.Hashimoto, H.Iyatomi and M.Tanaka, "Internet-based program for automatic discrimination of dermoscopic images between melanoma and Clark nevi," British Journal of Dermatology, No.150, 5, p.1041, 2004.

[18] H.Iyatomi, H.Oka, M.Tanaka et.al, "Quantitative assessment of tumour area extraction from dermoscopy images and evaluation of the computer-based methods for automatic melanoma diagnostic system," Journal of Melanoma Research, Vol.16, No.2, pp.183-190, 2006.

[19] M.Mastrolonardo, E.Conte and J.P.Zbilut, "A fractal analysis of skin pigmented lesions using the novel tool of the variogram technique," Chaos solutions and fractals, Vol.28, No.5, pp.1119-1135, 2006.

[20] S.Seidenari, G.Pellacani and C.Grana, "Asymmetry in dermoscopic melanocytic lesion images: a computer description based on colour distribution," Acta dermato-venereologica, Vol.86, No.2, pp.123-128, 2006.

[21] M.E. Celebi, H.A. Kingravi, Y.A. Aslandogan, and W.V. Stoecker, "Detection of Blue-White Veil Areas in Dermoscopy Images Using Machine Learning Techniques," Proc. of the SPIE Medical Imaging 2006 Conference, San Diego, CA, February 2006, 6144: 1861-1868.

[22] B.S.Everitt and G.Dunn, Applied Multivariate Data Analysis, London: Edward Arnold, pp.219-220, 1991.