International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 9, Number 24 (2014) pp. 27719-27731

http://www.ripublication.com

Detection and Classification of Skin Lesion in Dermoscopy images

Pranita P. Patil, Prof. Sonali A. Patil, Dr. V. R. Udupi

M.E. Student, K. J. Somaiya College of Engineering, Mumbai University, India.
E-mail: pranita.p@somaiya.edu
PhD from Shivaji University, Kolhapur and
Associate Professor in K. J. Somaiya college of engineering, IT department,
Mumbai University, India. E-mail: sonapatil_123@yahoo.co.in
working in department of Electronics and Communication Engineering,
GIT, Belgaum, India. Email: vishwa_u@yahoo.com

Abstract

Dermoscopy is used for the examination of the skin lesions using dermoscopy tool. Dermoscopy is mainly used to evaluate pigmented skin lesions. In this paper Detection and classification system is presented which accepts dermoscopic image containing lesion as input and predicts the class of lesion as benign or malignant. Pre-processing is applied on dermoscopic images and segmentation is performed using semi-automated thresholding, automated thresholding and k-means clustering based segmentation. Features are extracted from segmented image and given as an input to classifier. Various classifiers used are k-nearest neighbour, support vector machine and feed forward back propagation neural network. This paper also compares performances of classifiers used.

Index Terms— Thresholding, K-means, ABCD Rule, Support Vector Machine, K- nearest neighbour, feed forward back propagation, neural network, radial basis function.

1 INTRODUCTION

Skin lesion is a superficial growth of the skin that doesn't resemble the area surrounding it[1]. Skin lesion can be grouped as benign or malignant. Benign skin lesions are non-cancerous whereas malignant skin lesions are cancerous in nature. Common types of skin lesions taken into consideration are melanocytic nevi, Seborrhoeic keratosis, Basal cell carcinoma and Melanoma. Out of the above mentioned four types Melanocytic nevi and Seborrhoeic keratosis belong to class benign whereas melanoma and basal cell carcinoma belong to class malignant. Among malignant skin lesions, Melanoma is one of the deadliest form of cancer. It often resembles moles and some of them develop from moles. If melanoma is detected and treated in early stage then it is almost curable, whereas if it is not, then it may spread to other parts of body due to which there may be difficulty in treating the patient and it may lead to death [2]. Computer aided system would be beneficial in assisting in detection of Skin Lesions. Detection and Classification System can be used for the diagnosis in which dermoscopic image is taken as an input and the system should predict the lesion is benign or malignant. So to serve the purpose of detection of melanoma, CAD detection and classification system is proposed in this paper. This system would reduce unnecessary biopsy referrals.

The rest of the paper is organized as follows. Section 2 gives detailed literature survey. Section 3 describes the skin lesion detection and classification system. Section4 presents the experimental analysis of the system. Finally, Section 5 presents conclusion from the experimental results.

2 LITERATURE SURVEY

Lot of work is carried out in the field of bio-medical imagining. Dermoscopic image taken as an input needs to be pre-processed as there may be many artifacts like hair, air bubble present in that image. Different types of filters which can be used for pre-processing the image are Median Filter [3][4][13][16], prewitt filter[4], Contra-Harmonic Filter[11].

Next step is segmentation which enables us to separate the lesion from the surrounding skin. Various segmentation methods like thresholding, clustering and machine learning techniques are used for segmenting skin lesion. Jaleel [3], Madhankumar [4], Beuren [5] had used thresholding technique for segmentation. In [4] validation was performed against boundary tracing algorithm and in [5] segmented results were evaluated against ground truth segmentation. Other variations of thresholding techniques are adaptive thresholding [6], multi level thresholding [7] and fuzzy logic based approach [8]. In [9], Castillejos had presented three novel frameworks W-FCM, WCPSFCM, and WK-Means for skin lesion segmentation. In [10], R. Sowmya Devi had proposed and evaluated the performance of various fuzzy clustering techniques for skin lesion segmentation.

Feature extraction is also equally important as compared to that of segmentation in which features are extracted and given to classifier which is in charge of making decision whether it is benign or malignant lesion.

Different approaches for Feature extraction are ABCD Rule [4][16], Texture Analysis [13], histogram features [11] , Wavelet Packet Transform[16], 2D wavelet transform[3], GLCM[17].

Classification can be performed either by supervised learning or unsupervised learning algorithms. In [4], total dermoscopy score (TDS) was calculated and used for classifying benign lesions from melanoma. Dawei [11] and *Lucia Ballerini*[12] had used *K*-NN technique for classification. In [12] hierarchical approach for *K*-NN was used for classification purpose. The overall accuracy of this classifier was 74.3±2.5%. SVM classifier was used in [13][14]. In [13] Accuracy of the system was discussed in terms of sensitivity and specificity which was 80.76% and 85.57% respectively. Neural network was used for classification in[3][15] [16] [17][18]. In [3], Network training was performed using benign and malignant dataset features. Epochs of training were repeated until MSE was less than the goal. In [18] the best classification performance was achieved by a nonparametric 24-NN classifier. The overall performance of 61% was achieved by classifying into three classes and a performance of 73% for the class of melanomas. In a classification approach with two categories (malignant and dysplastic lesions against benign lesions), a sensitivity of 87% for the intervention class and a specificity of 92% was observed. In [19], Sadeghi had defined density ratio for automatic diagnosis system. The system was used for classification of moles and detection of skin cancer. The accuracy of the proposed system was 94.3%.

From literature survey we can conclude that different variations in thresholding were used for segmentation and proved to be an efficient technique for segmenting skin lesions. Hence semi-automated thresholding, automated thresholding and k-means based clustering were proposed for segmentation. Semi-automated thresholding is user dependent thresholding technique in which user needs to enter a threshold value whereas in automated thresholding the threshold value is updated iteratively so that it gives better segmentation results. Various classification techniques have been studied and based on that different variations in K-NN, SVM and Neural network classifier are proposed in this paper.

3 METHOD

This section describes the Detection and Classification system for skin lesion. Dermoscopic image is taken as an input. Pre-processing and Segmentation is applied which differentiates the skin lesion from surrounding healthy part of the skin. Features are extracted from the segmented lesion and classified based on those extracted features as benign or malignant.

3.1 Pre-processing

This section describes about the various pre-processing techniques used for skin lesion detection system. Pre-processing is important and crucial step. If pre-processing is done correctly then it results into better segmentation results.

In dermoscopic images there may be artifacts like hair present in it. It needs to be removed from skin for better segmentation results. Elimination of hair and hair like structures is needed for analyzing features more effectively as it can affect in the classification phase.

The median filter is a nonlinear digital filtering technique often used to remove spurious noise. It is used as pre-processing step which improves the image used for segmentation purpose. Median filtering ensures that unwanted structures are eliminated.

Contrast stretching is a simple image enhancement technique that attempts to improve the contrast in an image by stretching the range of intensity values it contains to span a desired range of values.

Steps followed in pre-processing dermoscopic image are as follows:
Step1: Accept input dermoscopic image.
Step2: Create structuring element of ‘disk’ shape with size=5.
Step3: Apply Bottom-Hat filtering.
Step4: Fill the region of interest.
Step5: Median filtering is applied.
Step6: Contrast Stretching is applied.

3.2 Segmentation

Segmentation is a process which subdivides image into multiple parts. The Level to subdivision is performed depends on the area of interest. As soon as the object of interest is obtained, segmentation process is stopped. In this paper, Semi-automated thresholding, automated thresholding and k-means clustering are proposed for skin lesion segmentation.

A. Semi-Automated Thresholding based Segmentation

Steps followed in Semi-Automated Segmentation are as follows:
Step1: Read Dermoscopic Image as input image.
Step2: Plot the histogram of the input image.
Step3: Based on the histogram, choose the threshold T.
Step4: Using the value of ‘T’, segment the image into object and background.
Step5: Image obtained from step4 is superimposed with original color image.

B. Automated Thresholding based Segmentation

Steps followed in Automated Segmentation are as follows:

Step1: Read Dermoscopic Image as input image.
Step2: Select an initial estimate for threshold, T.
Step3: Segment the image using T. This will produce two groups of pixels: G1, consisting of all pixels with intensity greater than T and G2, consisting of pixels with values less than or equal to T.
Step4: Compute the average intensity values for m1 and m2 for the pixels in regions G1 and G2, respectively.
Step5: Compute a new threshold value, T=(m1+m2)/2.
Step6: Repeat step3 through step5 until the difference in T in successive iterations is smaller than a predefined value, ΔT.
Step7: Segment the image using function im2bw(f,T/den).
Where den is an integer that scales the maximum value ratio T/den to 1, as required by function im2bw.
Step8: Result of step7 is superimposed with original dermoscopic image.

C. K means clustering based segmentation

Steps followed in K means clustering based Segmentation are as follows:
Step1: Read Dermoscopic Image as input image.
Step2: Resize the image to matrix 'I' of size 256 by 256.
Step3: Convert the matrix 'I' to one dimensional array.
Step4: Select any two Centroids.
Step5: Assign each pixel to the group that has the closest centroid.
Step6: When all objects have been assigned, recalculate the positions of both the centroids.
Step7: Repeat Steps 5 and 6 until the centroids no longer move.
Step 8: Resize the one dimensional matrix 'I' of size 256 by 256(matrix I contains index value where 1 represents background and 2 represents foreground).
Step9: Initialize a matrix 'S' of size 256 by 256 with all pixels set to 255.
Step10: If cluster index in matrix 'I' is equal to '1', then modify the corresponding pixel in matrix 'S' to '0' (When condition is satisfied pixel is set to background in matrix 'S').
Step10: matrix 'S' is superimposed with original dermoscopic image.

3.3 Feature Extraction

Feature extraction is representation and description of segmented image for further processing. Image is represented as n-dimensional feature space by extracting features and is further used for the classification process. 19 features were calculated which includes area, perimeter, circularity, border irregularity, major axis, minor axis, color variegation, mean intensities of RGB plane separately, entropy, energy, contrast, homogeneity, correlation, eccentricity and area difference.

3.5 Classification

Classification phase makes decision from extracted feature set. Set of extracted features from previous step are taken as input for training the different classifiers. Classification techniques proposed in this paper are k-nearest neighbor (*K*-NN), Support Vector Machine (SVM), Neural Network (NN) and are discussed as follows:

A. *K*-Nearest Neighbour

Steps followed in KNN based classification is as follows:
Step1: Accept Input set of features Xnew to be predicted.
Step2: Normalize the accepted values.
Step3: Find the no. of neighbouring points in the training set X that are nearest to Xnew.
Step4: Find the no. of neighbouring response values Y to those nearest points.
Step5: Assign the classification label Ynew that has largest posterior probability among the values in Y.

B. Support Vector Machine

Steps followed in SVM based classification is as follows:
Step1: Accept Input set of features to be predicted.
Step2: Normalize the accepted values.
Step3: Train an SVM Classifier.
Step4: Classifying New Data with an SVM Classifier.

C. Neural Network

Artificial Neural network refers to inter-connection between the neurons. Three parameters which define ANN are interconnection of nodes, updating weights by self learning process and an activation function that converts the weighted sum of interconnecting nodes into output function. Various supervised architectures of ANN are feed forward back propagation, radial basis. Feed forward is the basic and the simplest architecture and most commonly used. In feed forward network information flows in only direction with no feedback loops. Errors are propagated back during training. Feed Forward back propagation neural network is proposed in this paper and

Steps followed in feed forward back propagation neural network (FFBPNN) based classification is as follows:
Step1: Load input vectors and target vectors.
Step2: Create the network.
Step3: Set up the division of data.
Step4: Train the network.

Step5: Test the network.
Step6: Plot Confusion matrix.

4 EXPERIMENTAL RESULTS

This section represents the implementation of segmentation and classification techniques for skin lesion segmentation and classification system. All experiments are performed on Intel i5-460M processor with 4 GB memory using MATLAB R2013a.

Dataset of 110 images was considered. Dermoscopic images were downloaded from [20][21][22]. The proposed techniques have been experimented with four datasets i.e., melanocytic nevi, Seborrhoeic keratosis, Basal cell carcinoma and Melanoma. Each data set is divided into two parts, one is used for training and other is used for testing.

Figure 1 represents the GUI for CAD system for skin lesion detection and classification along with the selected image used for further processing.

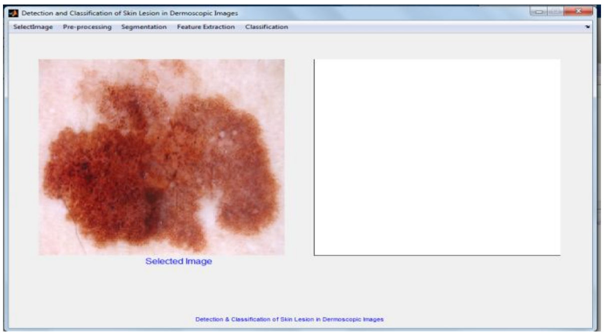

Figure 1: Selected Image

Techniques described in pre-processing are mandatory to be applied before segmentation step so as to enhance the selected image and obtain satisfactory results. Hair removal, median filtering and contrast stretching are applied for pre-processing as shown in figure 2, figure 3 and figure4 respectively.

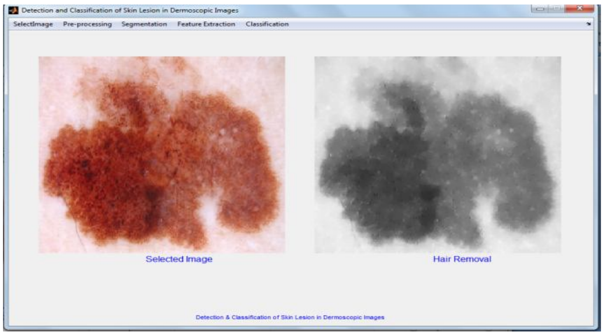

Figure 2: Image after Hair Removal

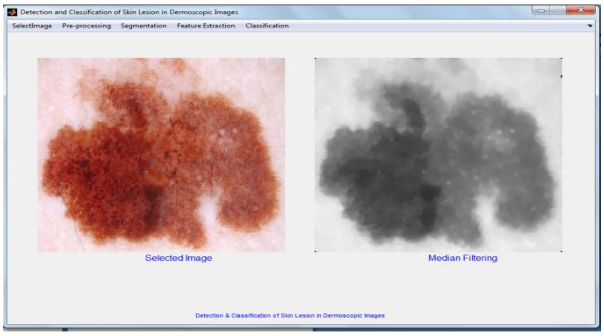

Figure 3: Median Filtering

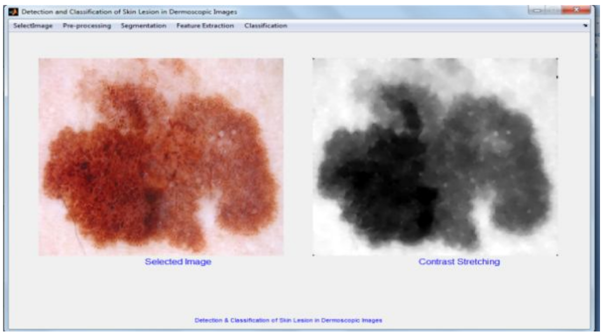

Figure 4: Contrast Stretching

Once pre-processing technique is performed next step is segmentation. Segmentation is also important step as it separates lesion from healthy skin. Figure 5 shows results of segmentation using Semi-Automated Thresholding. Results of semi-automated thresholding depend on the value of threshold. Figure 6 and Figure 7 shows results of segmentation using Automated Thresholding and K-means based segmentation respectively.

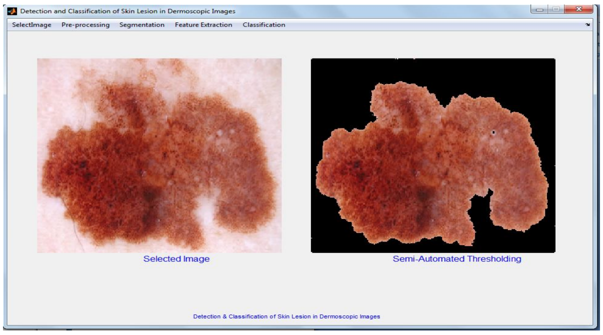

Figure 5: Semi-Automated Thresholding

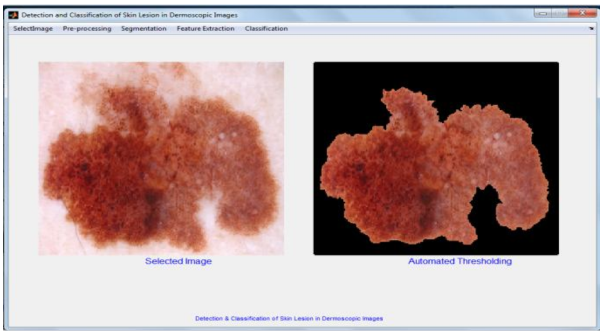

Figure 6: Automated Thresholding

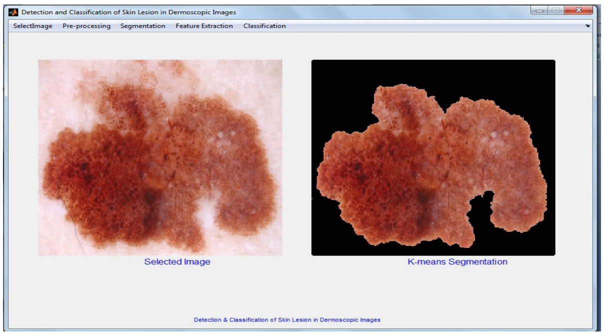

Figure 7: K-means Segmentation

Third step is feature extraction which extracts various features are as listed below in table 1.

Table1: Feature Extraction values

Features	Benign lesion	Malignant lesion
Area	5059	34428
Perimeter	363	1324
Circularity	0.4827	0.2469
Minor	79	207
Major	80	239
Border Irregularity	2.0719	4.0502
Cr	0.1353	0.3406
Cg	0.0763	0.1851
Cb	0.0524	0.1434
Mean R	7.5712	84.647
Mean G	3.5969	37.600
Mean B	2.1992	29.675
Entropy	0.9267	4.4973
Contrast	0.044	0.2787
Correlation	0.9103	0.9295
Energy	0.8534	0.2791
Homogeneity	0.9863	0.9246
Area Difference	93.286	4443.6
Eccentricity	0.1576	0.4998

Last step is to correctly classify the skin lesion as benign or malignant depending on the features extracted. Classification is performed using three techniques KNN, SVM and FFBPNN.

In KNN classification, various variations are performed using parameters nearest neighbour, break ties and different distance metrics. Table 2 describes the correctly classified results for KNN with parameter variation.

Table 2: KNN classification

Parameter	Type	Classified%	Overall%
Nearest Neighbour	Benign	92.45	89.09
	Malignant	85.96	
Break Ties	Benign	90.57	88.18
	Malignant	85.96	
Distance	Benign	90.57	89.09
	Malignant	87.72	

In SVM classification there are two types linear and non-linear. Kernel function value is varied. By setting kernel function value as linear we obtain linear SVM classifier and by setting value as rbf we get non linear SVM classifier. Table 3 describes the results of SVM classifier.

Table 3: SVM classification

Parameter	Type	Classified %	Overall %
Linear	Benign	90.57	88.18
	Malignant	85.96	
rbf	Benign	84.91	87.27
	Malignant	89.47	

In feed-forward back propagation neural network, hidden layers are varied from 5 to 30 and different training functions are used. Activation function used is tan-sigmoid. Table4 shows the results where number of hidden layers used are 10 and training function is set to 'trainlm' and overall performance obtained is 0.0843

Table 4: FFBPNN classification

Parameter	Type	Classified%	Overall %
FFBPNN	Benign	96.23	90.90
	Malignant	85.96	

5 CONCLUSION

In this paper Detection and Classification system was designed using Image Processing Toolbox from MATLAB. Segmentation algorithms like semi-automated thresholding, automated thresholding and k-means were tested on the dataset which produced satisfactory results. K-means correctly segments more images as compared to that of others. Feature extraction was performed in which various features were extracted and those features were given to various classifiers after normalization like FFBPNN, KNN, and SVM. Classification rate of KNN is 88.18±0.91 % whereas classification rate of SVM is 87.27±0.91 %. The overall accuracy achieved using FFBPNN is 90.90%. Different Variations in KNN was carried out in which nearest neighbour with k=1 was more accurate as compared to that of others. Linear SVM gave better performance as compared to that of svm with kernel value as 'rbf'.

REFERENCES

[1] Skin Lesion :http://www.healthline.com/health/skin-lesions.

[2] Skin cancer : www.skincancer.org

[3] Dr. J. Abdul Jaleel, Sibi Salim, Aswin.R.B, "Artificial Neural Network Based Detection of Skin Cancer", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 1, Issue 3, September 2012

[4] K. Madhankumar, P. Kumar, "Characterization of Skin Lesions", Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering , March 21-23, 2012

[5] Arlete Teresinha Beuren, Rodrigo Janasieivicz, Neusa Grando, Jacques Facon, "Skin Melanoma Segmentation by Morphological Approach", ICACCI'12, August 3-5, 2012, Chennai, T Nadu, India.

[6] Teresa Mendonca, Andre R. S. Marcal, Angela Vieira, Jacinto C. Nascimento, Margarida Silveira, Jorge S. Marques, Jorge Rozeira, "Comparison of Segmentation Methods for Automatic Diagnosis of Dermoscopy Images", Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, Lyon, France August 23-26, 2007.

[7] Jawad Humayun, Dr. Aamir Saeed Malik, Dr Nidal Kamel, "Multilevel Thresholding for segmentation of pigmented skin lesions", 978-1-61284-896-9/11/$26.00 ©2011 IEEE.

[8] M. Emin Y¨uksel, Senior Member, IEEE, and Murat Borlu, "Accurate Segmentation of Dermoscopic Images by Image Thresholding Based on Type-2 Fuzzy Logic", IEEE transactions on fuzzy systems, vol. 17, no. 4, august 2009

[9] Heydy Castillejos, Volodymyr Ponomaryov, Luis Nino-de-Rivera, Victor Golikov, "Wavelet TransformFuzzy Algorithms for Dermoscopic Image Segmentation", Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine Volume 2012.

[10] R. Sowmya Devi, Dr. L. Padma Suresh, Dr.K.L.Shunmuganathan, "Intelligent Fussy System Based Dermoscopic Image Segmentation for Melanoma Detection", Chennai and Dr.MGR University Second International Conference on Sustainable Energy and Intelligent System (SEISCON 2011).

[11] Dawei Nie, "Classification of melanoma and clark nevus skin lesions based on Medical Image Processing Techniques", 978-1-61284-840-2/11/$26.00 ©2011 IEEE.

[12] *Lucia Ballerini, Robert B. Fisher, Ben Aldridge, Jonathan Rees,* "Non-Melanoma Skin Lesion Classification Using Colour Image Data in a Hierarchical K-NN Classifier", 978-1-4577-1858-8/12/$26.00 ©2012 IEEE

[13] *Tarun Wadhawan, Ning Situ, Keith Lancaster, Xiaojing Yuan*, *George Zouridakis*, "SkinScan_c : A PORTABLE LIBRARY FOR MELANOMA DETECTION ON HANDHELD DEVICES", ISBI 2011.

[14] Yanal Wazaefi1, Sébastien Paris, Bernard Fertil, "Contribution of a classifier of skin lesions to the dermatologist's decision", 978-1-4673-2584-4/12/$31.00 ©2012 IEEE

[15] Suhail M. Odeh, "Using an Adaptive Neuro-Fuzzy Inference System (AnFis) Algorithm for Automatic Diagnosis of Skin Cancer", Journal of Communication and Computer 8 (2011) 751-755

[16] Andy Chiem, Adel Al-Jumaily, and Rami N. Khushaba, "Novel Hybrid System for Skin Lesion Detection", ISSNIP 2007

[17] Nabanita Bhattacharjee, Ranjan Parekh , "Skin Texture Analysis for Medical Diagnosis", *ICCCS'11*, February 12–14, 2011, Rourkela, Odisha, India.

[18] Harald Ganster, Axel Pinz, Reinhard Rohrer, Ernst Wildling, Michael Binder, and Harald Kittler, "Automated Melanoma Recognition", IEEE transactions on medical imaging, vol. 20, no. 3, march 2001.

[19] Maryam Sadeghi, Majid Razmara, Tim K. Lee, M.Stella Atkins, "A novel method for detection of pigment network in dermoscopic images using graphs", Computerized Medical Imaging and Graphics 35 (2011) 137–143

[20] http://www.dermnetnz.org

[21] http://www.pcds.org.uk

[22] www.dermoscopy.wordpress.com