Skin Research and Technology 2009; ■■: 1–11
Printed in Singapore · All rights reserved
doi: 10.1111/j.1600-0846.2009.00408.x

Skin Research and Technology

Pre-diagnostic digital imaging prediction model to discriminate between malignant melanoma and benign pigmented skin lesion

Jeppe H. Christensen[1], Mads B. T. Soerensen[1], Zhong Linghui[2], Sun Chen[2] and Morten O. Jensen[1,3]

[1]*Department of Electrical Engineering, Engineering College of Aarhus, Aarhus, Denmark,* [2]*Department of Computer and Information Technology, Shandong Jiaotong University, Jinan, China, and* [3]*Department of Biomedical Engineering, Engineering College of Aarhus, Aarhus, Denmark*

Background: Malignant cutaneous melanoma is the most deadly form of skin cancer with an increasing incidence over the past decades. The final diagnosis provided is typically based on a biopsy of the skin lesion under consideration. To assist the naked-eye examination and decision on whether or not a biopsy is necessary, digital image processing techniques provide promising results.

Hypothesis and aims: The hypothesis of this study was that a computer-aided assessment tool could assist the evaluation of a pigmented skin lesion. Hence, the overall aim was to discriminate between malignant and benign pigmented skin lesions using digital image processing.

Methods: Discriminating algorithms utilizing novel well-established morphological operations and methods were constructed. The algorithms were implemented utilizing graphical programming (LabVIEW Vision). Verification was performed with reference to an image database consisting of 97 pigmented skin lesion pictures of various resolutions and light distributions. The outcome of the algorithms was analysed statistically with MATLAB and a prediction model was constructed.

Results/Conclusion: The prediction model evaluates pigmented skin lesions with regards to the overall shape, border and colour distribution with a total of nine different discriminating parameters. The prediction model outputs an index score, and by using the optimal threshold value, a diagnostic accuracy of 77% in discriminating between malignant and benign skin lesions was obtained. This is an improvement compared with the naked-eye analysis performed by professionals, rendering the system a significant assistance in detecting malignant cutaneous melanoma.

Key words: malignant melanoma – receiver operating characteristics – automated vision analysis – LabVIEW – prediction model

Accepted for publication 5 August 2009

MALIGNANT CUTANEOUS melanoma is the most deadly form of skin cancer and the incidence has been alarmingly increasing over the past 30 years (1, 2). Despite increased public awareness and education including free cancer screenings, the incidence and mortality of malignant melanoma continue to rise (3).

Figure 1 shows the incidence rate regarding malignant melanoma for men and for women in a period of 33 years from 1970 to 2003 in Scandinavia.

The tendency is clear with a steady rise in incidence from 1970 to 2003. This can partly be explained by the greater awareness of malignant melanoma among people and professionals. This, however, states the importance of developing new and improved methods for diagnosis of malignant cutaneous melanoma.

The prognoses for the disease highly depend on which phase the diagnoses are given. If it is treated in the curable early stages, the prognosis is good (4).

A final and exact diagnosis can only be given subsequent to a biopsy of the inspected skin lesion. It is, therefore, a general practitioner's or a dermatologist's subjective judgment whether or not this biopsy is necessary (5). For this evaluation, the doctor will have a number of criteria available which all rely on a naked-eye analysis (1, 2, 4). In addition, remote analysis of skin lesions is an increasing topic in telemedicine.

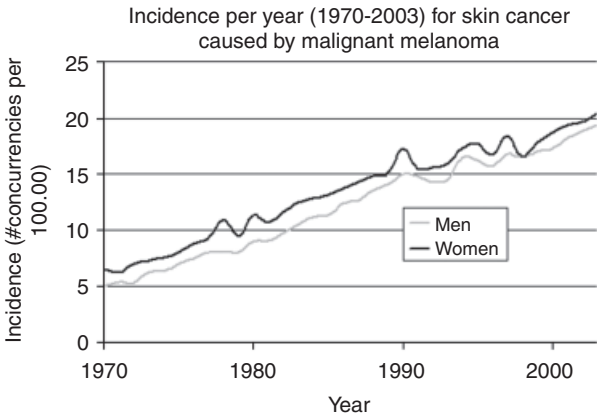

Fig. 1. Incidence of malignant melanoma from 1970 to 2003 in Scandinavia. The incidence is given in number of concurrences per 100,000.

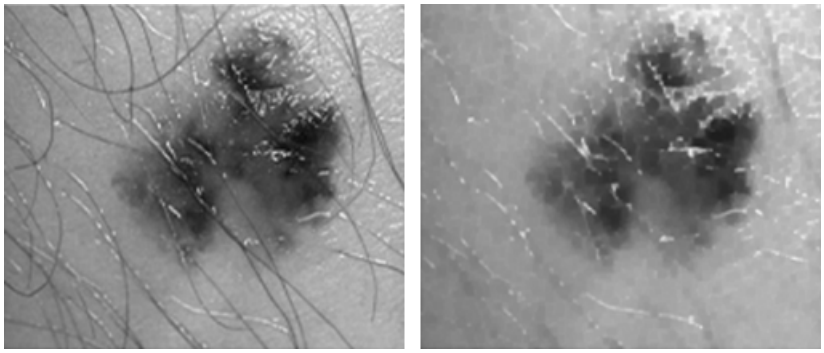

Fig. 2. Removing hairs morphologically. Image obtained from Online Atlas. Published online at http://www.dermis.net/doia/

In this work, a computer-assisted prediction model utilizing the most significant methods found in literature to visually discriminate between benign and malignant skin lesions is developed. The goal is to at least match the diagnostic accuracy of 64% achieved by dermatologists (1).

Image Pre-Processing

Segmentation partitions the digital image into two regions consisting of the skin lesion and the surrounding skin and thereby facilitates further analysis. An incorrect segmentation algorithm will result in an image analysis with a strong bias. Therefore, it is crucial that the segmentation algorithm is stabile and unaffected by the different appearances of skin lesions size, colour, contrast to the surrounding skin, etc. Furthermore, hair disturbs the segmentation of the skin lesion since their colours correlate strongly.

Removing artefacts

Using grey scale morphology, thin or thick hair can be removed from the image. This is done using a morphological closing operation, which is shown in the images in Fig. 2.

Threshold

Mendoca et al. (6) have recently shown that adaptive histogram thresholding gives the best result in segmentation of a skin lesion among methods such as watershed, neural networks, edge detection and others.

Adaptive threshold is furthermore considered as one of the simplest methods for segmentation and can therefore be implemented with a reduced amount of time use compared with other methods (6).

According to Ganster et al., (7) a threshold using the blue plane of the RGB colour model gave the best segmentation results. Furthermore, the blue plane represents the best contrast between the skin lesion and the surrounding skin (8).

Local adaptive threshold will compensate for non-uniform lighting changes throughout the image. This is done by adapting the threshold level individually to different segments of the image. The method selects an individual threshold for

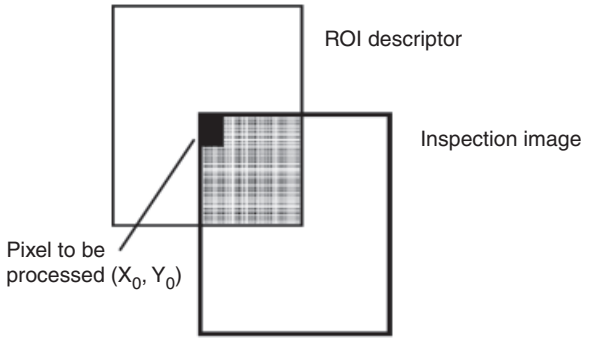

Fig. 3. The inspection image is processed with a mask of the same dimensions. Thereby the pixels to be processed will minimise use to one-fourth of the inspection image pixels.

each pixel based on the magnitude of the pixels in its local neighbourhood. Thereby, the effect of light variations throughout the image is minimized and the algorithm will be more stable with regard to input images from different sources.

Optimal threshold requires a region of interest (ROI) descriptor of a minimum of the same size as the skin lesion.

To overcome problems with different image and skin lesion sizes, the ROI descriptor is chosen to be the same size as the image under inspection. This ensures that a sufficient number of pixels are included in the calculation of the mean value within the ROI descriptor. The situation is depicted in Fig. 3.

The pixels included in the calculation of the mean value will be the intersection of the inspection image and the ROI descriptor (grey area in Fig. 3). Pixels of the ROI descriptor outside the inspection image will be 0.

The background corrected pixel (X_0,Y_0) is calculated as $I_B(X_0,Y_0) = I(X_0,Y_0) - m(X_0,Y_0)$, where $I(X_0,Y_0)$ is the original pixel value of the inspection image and $m(X_0,Y_0)$ is the local mean at pixel (X_0,Y_0) calculated on basis of the ROI descriptor.

After the background corrected image is created, an auto threshold is performed using Otsu's algorithm as described by Liao et al. (9).

Image post processing

Unwanted noise particles and particles touching the border of the image are removed by applying simple morphological operations.

Finally, the segmentation is done using a simple AND function with the original RGB colour image. Pixels within the border will be the skin lesion and pixels outside the border will be the surrounding skin.

The final result is shown in Fig. 4.

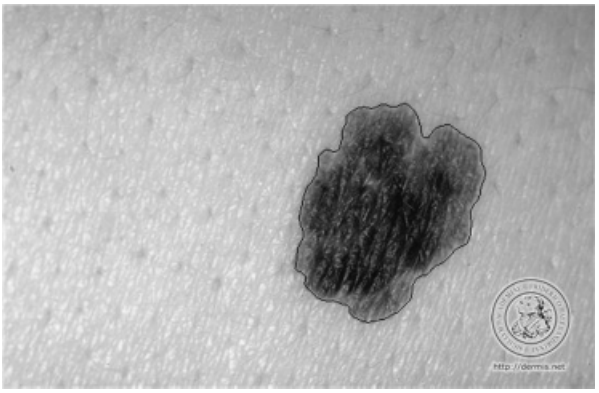

Fig. 4. Final segmentation result of the threshold algorithm. Diepgen TL, et al. Dermatology Online Atlas. Published online at: http://www.dermis.net/doia/

Feature Extraction

Several features have been used as classifiers in the prediction model.

Border parameter analysis

Border irregularities of a skin lesion are generally seen divided into two subgroups: structural irregularities and texture irregularities. Texture irregularities are small variations along the edge of the whole skin lesion (roughness). Both situations are depicted in Fig. 5.

The textural irregularities are highly sensitive to noise and are very dependent on precise segmentation which is why they are not included as a feature in the model (8).

Mean curvature (MC) vs. smooth

Melanomas have a higher tendency to exhibit protrusions and indentations, which can be detected by analysing the curvature function of a skin lesion border (10).

Curvature extremes define the tip points of concave or convex segments in the skin lesion. A protrusion is then defined as a segment beginning with a *concave* extreme followed by a *convex* extreme and ending with a *concave* extreme. Likewise, an indentation is defined as a segment beginning with a *convex* extreme followed by a *concave* extreme and ending with a *convex* extreme.

By taking the absolute value to the curvature function, the magnitude of the mean value will indicate the curvedness. The mean value will increase in magnitude when numerous protrusions and indentations are present in the border.

To compromise the variety in image quality, a relationship between the original border and a

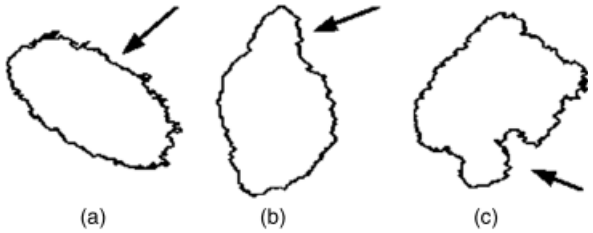

Fig. 5. Three skin lesions with border irregularities: (a) Textural irregularities. (b) Structural protrusion. (c) Structural indentation.

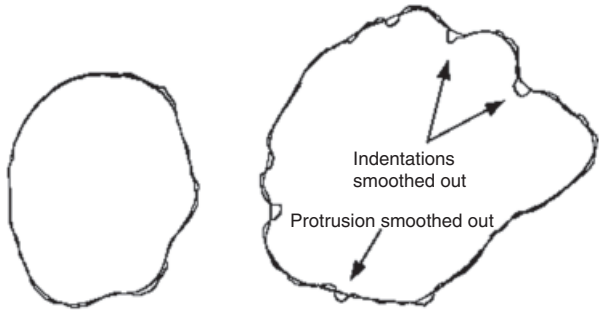

Fig. 6. The left figure shows the outline of a benign skin lesion and the corresponding smoothed version. To the right the outline of a malignant skin lesion is shown together with its smoothed version. Several indentations and protrusions are eliminated after smoothing the malignant skin lesion.

smooth version of the border is utilized. Figure 6 shows two skin lesions and their respective smoothed versions. The lesion to the left is benign, and the smoothed version can hardly be distinguished from the original skin lesion border. This will result in mean curvature values of almost equal magnitude. The lesion to the right is malignant and several indentations are smoothed out from the original lesion border. This will result in an evident difference in mean curvature values.

Global edge irregularity

The global edge irregularity of the skin lesion can be evaluated utilizing the border irregularity index (BI) mentioned by Farina et al. (11). BI is defined as

$$\mathrm{BI} = \frac{perim_{\mathrm{C}}}{perim_{\mathrm{L}}}$$

In the above definition, $perim_{\mathrm{C}}$ is the perimeter of the convex hull enclosing the lesion (see figure 7) and $perim_{\mathrm{L}}$ is the perimeter of the skin lesion. It is clearly seen by the definition of BI that an ideally round lesion with regular borders will result in a parameter score equal to 1. BI will then decrease proportional with the degree of irregularity of the lesion borders because of the increase in $perim_{\mathrm{L}}$. BI considers the skin lesion globally and is able to detect large indentations or protrusions, which would not have been identified by MC because of the small radius of curvature in this case.

Additionally indentations and protrusions can be measured using an area index (AI) which rely on a relationship between the area of the skin lesion and the area of its convex hull.

The AI is defined as

$$\mathrm{AI} = \frac{area_{\mathrm{C}}}{area_{\mathrm{L}}}$$

In the above definition, $area_{\mathrm{C}}$ is the area of the convex hull containing the lesion and $area_{\mathrm{L}}$ is the area of the skin lesion. In digital imaging, area is equal to the number of pixels.

It is seen that for a lesion with no protrusions and indentations, the AI will equal to 1. If protrusions or indentations exist within the skin lesion, $area_{\mathrm{C}}$ will increase more than $area_{\mathrm{L}}$ and the parameter score will be higher.

Figure 7 shows the profile of two skin lesions from the image database and their convex hull.

Best-fit ellipse index (BEI)

According to Chang et al., (12) melanomas have a greater tendency than benign skin lesions to have an ellipse-shaped form.

The shape of the skin lesion can be evaluated using a BEI. The BEI is defined as

$$\mathrm{BEI} = \frac{ab}{2\pi(a^2 + b^2)}$$

Where a and b are the length of the major and minor axes of the best-fit ellipse surrounding the skin lesion.

In earlier clinical studies, dermatologists examined 60 clinical image lesions and yielded two errors using the BEI (12).

Standard deviation of mean radius (SD)

Manousaki et al. (1) suggests the use of SD of the skin lesion as a geometric factor for evaluating

Fig. 7. Profile of two skin lesions and their surrounding convex hull. The skin lesion to the left is benign and the one to the right is malignant.

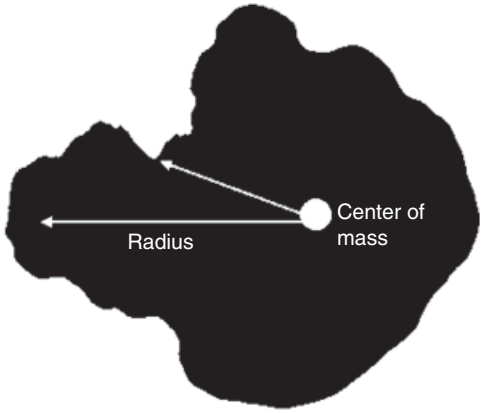

Fig. 8. Profile of a malignant skin lesion. The radius is calculated from the center of mass in the skin lesion.

the skin lesion border. The radius is calculated from the centre of mass in the skin lesion and with respect to a number of different points along the skin lesion perimeter. Next, a standard deviation is derived from these measurements. A large standard deviation will imply the presence of structural irregularities. The concept is depicted in Fig. 8.

In the above example, the protrusion on the left side of the skin lesion will cause the radius to deviate from the mean radius. This will again cause a larger SD than in the situation where the protrusions are not present.

Roundness (R)

According to Farina et al., (11) the R factor of a skin lesion is an important aspect to consider when looking for melanoma. The R factor tells how much the skin lesion resembles a circle. It is defined as

$$R = \frac{4\pi \cdot area_{\mathrm{L}}}{(perim_{\mathrm{L}})^2}$$

It is seen that when the R factor is calculated for a circle, R becomes

$$R = \frac{4\pi \cdot area_{\mathrm{L}}}{(perim_{\mathrm{L}})^2} = \frac{4\pi r^2 \pi}{(2r\pi)^2} = \frac{4\pi^2 r^2}{4\pi^2 r^2} = 1$$

Protrusions and indentations in a skin lesion will create a disproportion in area vs. perimeter. Because the skin lesion perimeter is squared, R will decrease as a result of protrusions and indentations.

Heywood circularity index (HCI)

In the same way that the roundness index, described above, compares the skin lesion with a circle, the HCI utilizes a comparison between the skin lesion and a circle. The HCI is defined as

$$\mathrm{HCI} = \frac{perim_{\mathrm{L}}}{2\sqrt{\pi area_{\mathrm{L}}}}$$

$perim_{\mathrm{L}}$ is the skin lesion perimeter and $area_{\mathrm{L}}$ is the skin lesion area. Thereby, the HCI is the perimeter (P) divided by a circle with the same area as the skin lesion. This is seen when the definition of the area of a circle is inserted in the formula

$$\mathrm{HCI} = \frac{P}{2\sqrt{\pi\pi r^2}} = \frac{P}{2\pi r} = \frac{P}{P}$$

It is then noticed that for a perfectly round circle, the HCI becomes 1. If the perimeter length P is increased and the area A is maintained due to structural irregularities, the HCI increases.

Colour parameter analysis

Before analysing the skin lesion with regards to its colours, the surrounding skin is removed from the image using the threshold algorithm described above. Furthermore, the border thickness is dilated. This ensures that no skin situated between the segmented skin lesion border and the actual skin lesion is included in the analysis. The border region of a skin lesion is often corrupted with light reflections or shadows causing a bias in the colour analysis.

To reduce the bias from these light reflections, all pixels with greyscale values $>70\%$ of the maximum value within the skin lesion is excluded from the colour analysis.

Furthermore, to create consensus between different image qualities, all pixel values are expanded to cover the full dynamic intensity range (grey level intensity from 0 to 255).

In this study, all implemented methods are calculated from the greyscale histogram of either the red or blue colour plane of the RGB colour model or

the luminance colour plane of the HSL colour model to identify the optimal colour plane to use.

The green colour plane of the RGB colour model contains little information with regard to the skin lesion because of the dominance in red and blue colour spectrum and will therefore not be evaluated (1).

Mean greyscale value (MGV)

According to Seidenari et al. (13) and Manousaki et al. (1) a difference in benign and malignant skin lesions can be measured by calculating the MGV for the skin lesion. The mean | x is the arithmetic average of the pixel values contained within the skin lesion, calculated as

$$\bar{x} = \frac{1}{N}\sum_{i=0}^{N} p_i$$

p_i is the i-th pixel intensity and N is the total amount of pixels within the skin lesion.

Standard deviation greyscale histogram (SDG)

The SDG is a measure of the dispersion of pixel values from the mean grey scale value. If the grey scale pixel values are concentrated close to the mean value in the image, the SDG is small.

According to Seidenari et al. (13) and Manousaki et al., (1) melanomas exhibit a heterogeneous distribution of grey scale values within the lesion. This will result in a higher SDG.

The standard deviation, σ, is calculated as the square root of the second standardized moment

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} (p_i - \overline{x})^2}$$

p_i is the i-th pixel intensity, x is the mean intensity and N is the total amount of pixels within the skin lesion.

The above calculation of variance is based on Gaussian distributions. This is not necessarily the distribution when analyzing skin lesion images. However SDG is still a valid measure despite other histogram distributions.

Histogram waveform analysis

According to Manousaki et al., (1) the irregularity in colour distribution within the lesion is an indicator of malignancy. It is therefore relevant to examine the shape of the greyscale histogram produced by the red, blue or luminance colour plane of the skin lesion.

Fig. 9. Different distribution functions with same standard deviation but different shape.

To measure the degree of irregularity consisted within a histogram waveform, the number of peaks exceeding a certain threshold can be counted (Peaks). Perfectly smooth histogram waveforms will have no peaks resulting from irregularity.

Kurtosis

The use of SDG will not always be a descriptor of the overall waveform shape which can be seen by observing figure 9. The mean values and the standard deviations of the two histograms | in Fig. 9 is equal, even though the shape of the histogram to the right seem more peaked. Malignant skin lesions tend to have more flat distributions as the one shown to the left in the figure, resulting from higher deviation of greyscale values. A measurement for these tendencies is kurtosis. It measures if the pixel distribution are peaked or flat relative to a normal distribution. Histograms with a higher kurtosis tend to have a distinct peak near the mean and then decline rather hastily, and have fatter tails. Histograms with lower kurtosis tend to have a flat top near the mean and have wider shoulders.

Kurtosis, σ^4, is the fourth standardized moment and is calculated as

$$\sigma^4 = \frac{1}{N}\sum_{i=0}^{N-1} (p_i - \overline{x})^4$$

p_i is the i-th pixel intensity, x is the mean intensity and N is the total amount of pixels within the skin lesion.

Prediction Model

A prediction model was developed to discriminate between malignant and benign skin lesions.

The procedure of developing this prediction model involves two steps:

The first step is to individually evaluate the parameters described above, of their significance to the final prediction model. This is to avoid including parameters with no or negative association to the final outcome.

The second step is to join the included parameters and by this, develop a final prediction model.

Statistical data analysis

A total of 97 skin lesions (43 benign and 54 malignant) were successfully segmented and analysed with respect to the above described parameters.

Evaluation of border parameters

The evaluation of border parameters regarding their ability to function as classifiers is done by calculating receiver operating characteristics (ROC) curves, by which the sensitivity and specificity of the prediction model can be evaluated. One ROC curve is calculated for each parameter.

To compare the different parameters against each other, the area under the ROC curve (AUC) and the optimal accuracy (ACC) is calculated.

Table 1 shows AUC and ACC for all the border parameters.

It is seen that parameters MC and SD exhibit the lowest AUC and ACC score. The objective of this project is to reach a discriminating accuracy better than 64% in the final prediction model.

The two parameters MC and SD both have an accuracy of approximately 64%. Furthermore, they have the lowest AUC score, which indicates no significant impact in the final prediction model. Therefore, these parameters are excluded as classifiers in the final prediction model.

Evaluation of colour parameters

Colour parameter ROC curves are calculated for each of the three selected colour planes, red, blue and luminance. They are then compared in one plot to identify which colour plane is optimal with regards to discrimination.

TABLE 1. Parameter AUC and ACC score.

Parameter	AUC	ACC
MC	0.61	0.63
BI	0.68	0.68
AI	0.73	0.69
BEI	0.71	0.70
SD	0.60	0.65
HCI	0.71	0.70
R	0.71	0.71

AUC, ACC, area under the ROC curve (AUC) and the optimal accuracy (ACC); MC, mean curvature; BI, border irregularity index; AI, area index; BEI, best-fit ellipse index; SD, standard deviation of mean radius; HCI, Heywood circularity index; R, roundness.

TABLE 2. Parameter AUC and ACC score

Parameter	AUC	ACC	Colour plane
MGV	0.72	0.68	Red
SDC	0.84	0.78	Red
Peaks	0.75	0.73	Red
Kurtosis	0.66	0.66	Blue

AUC, ACC, area under the ROC curve (AUC) and the optimal accuracy (ACC); MGV, mean greyscale value; SDC, standard deviation greyscale histogram.

Additionally, AUC and ACC are calculated. The optimal result and the corresponding colour plane are listed in Table 2.

As seen in the table, all of the colour parameters, ACC are above 64%. Furthermore, they all exhibit relatively good AUC. The kurtosis parameter gave the highest AUC score using the blue colour plane.

Building a classification model

To reach a final classification of the skin lesion, all measurements of benign and malignant skin lesions, respectively, are summed together. By doing this, the importance of every parameter can be examined and one single parameter can be identified to discriminate between skin lesions.

The summing will consist of two cases:

Benign skin lesions are expected to exhibit a *higher* parameter score than malignant skin lesions.

Benign skin lesions are expected to exhibit a *lower* parameter score than malignant skin lesions.

The two cases are summed together using opposing signs, thereby, creating a consensus between all the parameters.

Before summing the data arrays from all the parameters, the difference in the parameter score must be considered. This is to ensure that a parameter exhibiting a score with a low magnitude also has an influence on the sum compared with a parameter score with a high magnitude.

The levelling of dynamic range is done by multiplying each parameter data array with the factor X defined as

$$X = 100/\frac{1}{9}\sum_{i=0}^{8} B_i$$

where B_i is the parameter measurement for the benign skin lesions. Hence all parameter scores are lifted or lowered so that the benign skin

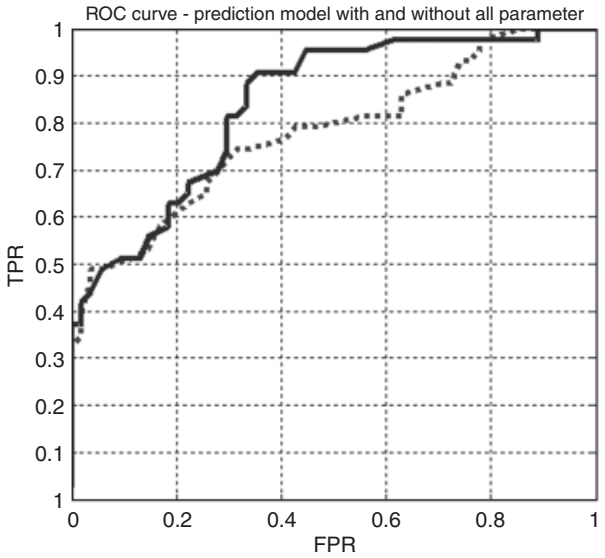

Fig. 10. Receiver operating characteristics curve showing characteristics before and after excluding of parameter MC and SD. The dotted line represents the ROC before excluding parameters and the solid line represents the ROC after excluding the parameters.

lesions mean value is 100. It is important to mention that the factor *X* is found for the data arrays containing data from benign skin lesions only but is multiplied on both benign and malignant skin lesion data.

After summing the scaled parameters together, a ROC curve is calculated (see figure 8). The true-positive rate (TPR) on the ROC curve (y-axis) is equal to the sensitivity of the prediction model. The false-positive rate (FPR) is equal to 1- specificity of the prediction model (x-axis). Every point plotted on the ROC curve is a possible threshold value and thereby a possible discrimination line. The solid curve shows the prediction model after the parameters MC and SD are excluded.

It is clearly seen on the ROC curve that the prediction model will perform optimally with the parameters excluded.

Furthermore, AUC was calculated to 0.78 with all parameters included and 0.84 with MC and SD excluded. This indicates an improvement of the prediction model with 7.7%.

Weighting of parameters

All results presented above rely on an equal weighting of the parameters used. As it is suggested by Voigt and Classens (4), the individual parameters needs to be distinctly weighted. This will increase the final diagnostic accuracy of the prediction model because not all parameters are equally important to the final prediction outcome.

To assign weighting scales to the individual parameters, a dedicated program has been developed. The program utilizes a visual assessment of the influence of weighting scales on the individual parameters. The first step in the optimization of the prediction model is to exclude the parameters which are not to be used. Next, parameter SDG, AI, Peaks and *R* are negated to build a consensus between all parameters.

To assist the visual assessment of the influence of changing the parameter weighting scales, a ROC curve and the corresponding AUC are calculated continuously. The objective is to change the parameter weighting scales so that AUC of the prediction model is maximized.

The highest possible AUC calculated was 0.87. The impact of using parameter weighting scales thereby results in an increase in AUC of 3.6%.

In Fig. 11, the ROC curve are shown with and without the parameter weighting scales.

Discrimination value

The optimal discrimination value between malignant and benign skin lesions is found by examining the ACC of the prediction model.

The ACC is equal to the total number of correctly diagnosed instances in relation to the

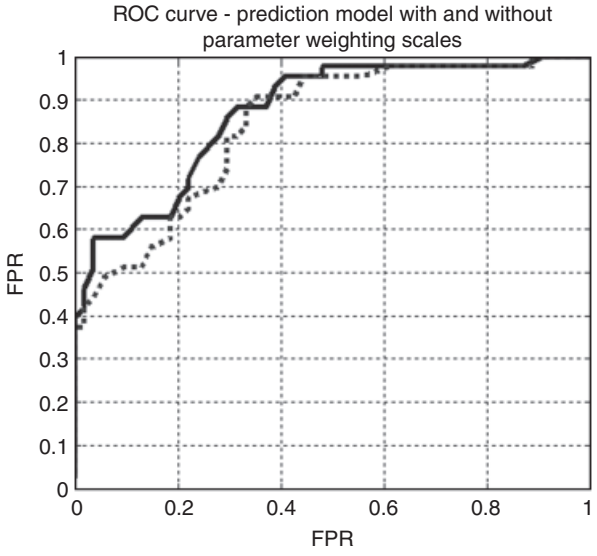

Fig. 11. Receiver operating characteristics curves showing the influence of weighting scales. The dotted curve is without the use of weighting scale and the solid curve is with the use of weighting scales.

total number of instances. It is calculated as a function of the amount of correctly diagnosed positive (benign) instances, true-positive (TP) and the amount of correctly diagnosed negative (malignant) instances, true-negative (TN)

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{P + N}$$

P and N are the total amount of positive (benign) and negative (malignant) instances, respectively.

By computing ACC for a number of threshold levels in the dynamic range of the prediction model score, a function of the accuracy vs. threshold level is produced. The highest accuracy is found at a threshold level of -43.

Using ROC curve analyse

To acquire a balanced relationship between sensitivity and specificity, both are plotted in a three-axis graph showing sensitivity and specificity vs. threshold level in the dynamic range of the prediction model score. This is seen in Fig. 12.

From the graph, it is seen that the optimal threshold should be -50.3 - the intersection between sensitivity and specificity. The plotted data around the threshold are shown in Table 3.

From the table, it is seen that by choosing a threshold level of -50.3, we can obtain a sensitivity of 76.74% and a specificity of 75.93%. If we choose a threshold level of -35.2 in a prediction model score instead, we would achieve a sensitivity of 86.05%. This is an improvement in sensitivity of 9.31 percentage point (12.1%) at the expense of a 5.56 percentage point (7.9%) lack in specificity.

Results

The prediction model developed has an outcome score in the range -286 to 215 for the 97 images used. The prediction model is developed so that benign skin lesion scores are higher than malignant skin lesion scores.

The final prediction model results in a ROC shown in Fig. 13.

The optimal threshold is in the range -35 to -40. By choosing a threshold value of -35.2, a sensitivity of 86% and a specificity of 70.37% is achieved.

A final contingency table is calculated using the above stated threshold. The data are shown in Table 4.

Next, a final accuracy for the developed prediction model using a threshold of -35.2 is calculated as ACC = 0.77.

Discussion

The objective was to reach an accuracy of at least 64%. This accuracy corresponds to that an experienced dermatologist would achieve (1). The final prediction model developed exhibits an accuracy of 77% and will thereby represent an improvement in diagnostic accuracy.

In spite of the many different appearances and qualities of the images in the image database, it has been possible to develop a prediction model with an accuracy better than that expected from an experienced dermatologist. The algorithms are thereby stable and not dependent of the quality of the image under inspection. These new tools can also be utilized in telemedicine, where images can be uploaded to an automated web-based database and analyzed.

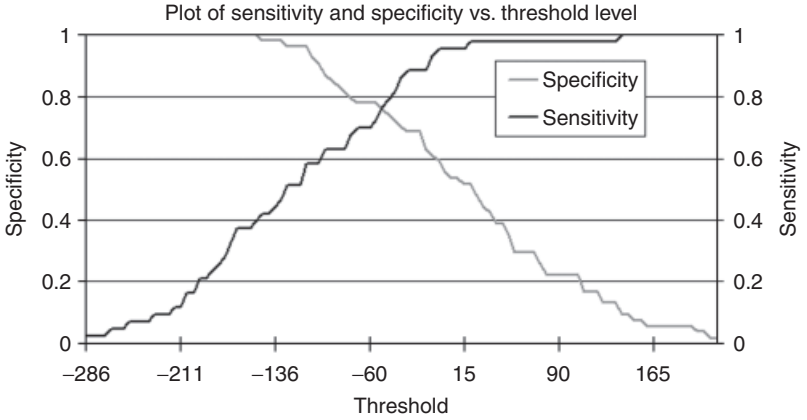

Fig. 12. Plot of sensitivity and specificity vs. threshold value. The optimal balance is identified in the intersection of sensitivity with specificity.

TABLE 3. Sensitivity and specificity of the prediction model at different threshold values.

Threshold	Sensitivity (%)	Specificity (%)
−65.3	69.77	77.78
−60.3	69.77	77.78
−55.3	72.09	77.78
−50.3	76.74	75.93
−45.2	79.07	74.07
−40.2	81.40	72.22
−35.2	86.05	70.37

TABLE 4. Contingency for the final prediction model.

	Actual diagnosis		
	Malignant	Benign	Total
Prediction model diagnosis			
Malignant	TP = 37	FP = 16	53
Benign	FN = 6	TN = 38	44
Total	43	54	

TP, true-positive; TN, true-negative; FP, false-positive; FN, false-negative.

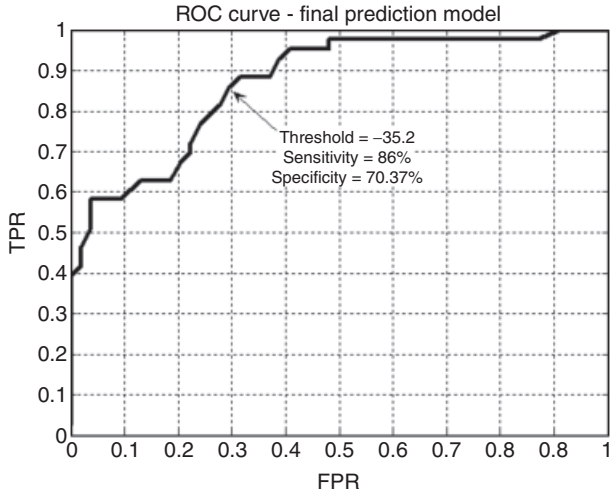

Fig. 13. Final receiver operating characteristics curve developed from the results from the prediction model.

The parameter MC was not implemented because of a low AUC score (0.61). This implies that the parameter is poor in discriminating between the two types of skin lesions. The reason for the low AUC score could be identified in the segmentation algorithm used. The spatial resolution of the inspection image is a highly determining factor in how well and precise the borders of the skin lesion is represented and segmented. The image database contains a wide range of different spatial resolution images. The curvature function in a low-resolution image will not be as detailed as in a high-resolution image and the mean curvature value will be highly dependent of this resolution. The parameter is therefore a poor discriminator when working with images of different resolution.

The parameter SD, like MC, exhibits an unsatisfactory AUC score (0.60). The SD is calculated as an absolute value, in spite of the fact that spatial resolutions vary in the images. An image where the skin lesion makes up half the amount of pixels as opposed to a second inspection image, will exhibit lower SD even though their shape are alike. To solve this problem, the SD needs to be related to the actual skin lesion size.

Three of the parameters for evaluating the colour behaviour of the skin lesion exhibit the best results using the red colour plane of the RGB colour model. The parameter Kurtosis exhibits the best results using the blue colour plane of the RGB colour model. This indicates that optimal result prediction can be achieved using a combination of different colour planes.

The results presented with regard to the final ROC, are based on skin lesions from the image database. It is seen that the amount of false-positive (FP) diagnoses is higher than the amount of false-negative (FN) diagnoses. It is a difficult choice when dealing with medical diagnosis whether it is better to diagnose a positive instance as negative or a negative instance as positive. With regards to this study, the FP rate should be higher than the FN rate, creating a safety margin. It is better to be treated for a malignant melanoma even though it actually is benign, than not receive any treatment at all while suffering from a malignant skin lesion.

Conclusion

The work presented in this paper shows that it is possible to develop and implement a pre-diagnostic evaluation tool for skin lesions using LabVIEW. The evaluation tool is meant to help clinicians and dermatologist in the diagnosis of skin lesions, especially in distinguishing between malignant melanomas and benign melanocytic skin lesions. The evaluation tool consists of a prediction model, which on the basis of the skin lesion border, shape and colour behaviour will predict a diagnosis.

The chosen discrimination value of a prediction model score of −35.2 results in a

sensitivity of 86% and a specificity of 70.37% and is therefore concluded to be the optimal choice.

The prediction model is built by statistical analysis of nine different evaluating parameters, which evaluate the inspected skin lesion with regards to its overall shape, border irregularity and grey scale histogram behaviour. All nine parameters contribute with different significances to the prediction model outcome.

The final accuracy achieved is based upon an analysis of a total of 97 skin lesion images containing 43 benign skin lesions and 54 melanomas. These images have also been used in developing the discriminating algorithms, and the final accuracy achieved needs therefore be verified by an analysis of more images. This analysis is also necessary before making use of the prediction model.

Using digital image processing technique as an alternative to the naked-eye analysis performed by a dermatologist has been proven possible. It should by no means replace an experienced dermatologist, but could be used as guidance to determine whether or not a certain skin lesion should be examined by a professional.

References

1. Manousaki AG, Manios AG, Tsompanaki EI, Panayiotides JG, Tsiftsis DD, Kostaki AK, Tosca AD. A simple digital image processing system to aid in melanoma diagnosis in an everyday melanocytic skin lesion unit. A preliminary report. Int J Dermatol 2006; 45: 402–410.
2. Calonje E. The histological reporting of melanoma. J Clin Pathol 2000; 53: 587–590.
3. Lang PG. Current concepts in the management of patients with melanoma. Am J Clin Dermatol 2002; 3: 401–426.
4. Voigt H., Classsen R. Computer vision and digital imaging technology in melanoma detection. Semin Oncol 2002; 4: 308–327.
5. Available at http://www.cancer.dk, visited on 10/9, 2007.
6. Mendonca T, Marcal ARS, Vieira A, Nascimento JC, Silveira M, Marques JS, Rozeira J. 2007. Comparison of segmentation methods for automatic diagnosis of dermoscopy images. *Conference of the IEEE EMBS*, Cité Internationale, Lyon, France, August 23–26.
7. Ganster H, Pinz A, Rodrer R, Wildling E, Binder M, Kittler H. Automated melanoma recognition. IEEE Trans Med Imaging 2001; 20: 233–239.
8. Lee TK, Claridge E. Predictive power of irregular border shapes for malignant melanomas. Skin Res Technol 2005; 11: 1–8.
9. Liao P, Chen T, Chung P. A fast algorithm for multilevel thresholding. J Inform Sci Eng 2001; 17: 713–727.
10. Lee TK, McLean DI, Stella AM. Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions. Med Image Anal 2003; 7: 47–64.
11. Farina B, Bartoli C, Bono A, Colombo A, Lualdi M, Tragni G, Marchesini R. Multispectral imaging approach in the diagnosis of cutaneous melanoma: potentiality and limits. Phys Med Biol 2000; 45: 1243–1254.
12. Chang R, Stanley J, Moss RH, Stoecker WV. A systematic heuristic approach for feature selection for melanoma discrimination using clinical images. Skin Res Technol 2005; 11: 165–178.
13. Seidenari S, Pellacani G, Grana C. Pigment distribution in melanocytic lesion images: a digital parameter to be employed for computer-aided diagnosis. Skin Res Technol 2005; 11: 236–241.

Address:
Jeppe Hoey Christensen
Strandvejen 8b, 3. Tv
8000 Aarhus
Denmark
Tel: +45 28760559
e-mail: jabbahc@hotmail.com