Automated Classification of Common Skin Lesions Using Bioinspired Features

D. Csabai,[1] K. Szalai,[1,2] M. Gyöngy[1]

[1]Ultrasound Laboratory, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary.
[2]Department of Dermatology, Venereology and Dermatooncology, Semmelweis University, Budapest, Hungary.

Abstract— **Ultrasound (US) imaging of skin lesions provides information supplementary to dermoscopy and helps in improving diagnostic accuracy. The aim of the current work is to explore the feasibility of using ultrasound image features derived from radiological experience to distinguish between common skin lesions. 5-18 MHz B-mode ultrasound images were acquired of incoming patients. Images containing lesions 1-2 mm thick were selected (N=248), with histology used to diagnose suspicious lesions. 73 melanomas, 130 BCC, and 45 nevi were studied. Following semi-automatic segmentation, a number of relevant features expressing the geometry of the lesion boundary and boundary layer, as well as the image characteristics of the lesion, lesion boundary layer, and post-lesion region were considered. With the exception of lesion echogenicity, all features had an area under the curve (AUC) value of above 0.70. The AdaBoost and Support Vector Machine (SVM) classifiers were then trained and tested using cross-validation of 50 random equal populations of melanomas, BCC, and nevi; each population was then 2-fold (holdout) cross-validated 50 times. When detecting one group against two other groups, the detection of cancerous lesions fared best, with an AUC of at least 0.84 and a specificity of at least 19% at 100% sensitivity for both classifiers. The results demonstrate the potential of clinically useful ultrasound-based automatic differential diagnosis of skin lesions, which could perhaps be attained by better segmentation, having more training data, using several images of the same lesion when performing classification, as well as refinements in the definition of image features.**

Keywords-skin, dermatology, computer-assisted diagnosis, melanoma

I. INTRODUCTION

Skin cancer is the most common type of cancer in the USA, with about 9000 people dying every year from its most dangerous form, melanoma [1]. Only 1 percent of skin cancer cases in the USA turn out to be melanoma [2]; the total costs of non-melanoma and melanoma skin cancer treatment are $4.8 billion and $3.3 billion, respectively [3]. As with most cancers, early diagnosis is essential: while the 5-year survival rate for localized melanoma is 98%, this drops to 63% for regional spread of the disease, and 17% for distant metastasis. [2]

There are many strategies for early skin cancer detection. These are now briefly presented from the context of melanoma detection, with the current paragraph being based on the review by Rigel et al. [4]. Most melanoma detection methods rely on the superficial examination of the skin, based on the ABCDE mnemonic for finding "Asymmetry, Border irregularity, Color variegation, Diameter >6 mm", supplemented by the recent addition of a temporally "Evolving" lesion [4]. These symptoms may be checked by oneself, one's partner, or a doctor. Doctors are generally better able to detect thinner melanomas. Doctor diagnosis is improved when they are a dermatologist with at least 5 years of experience, and when they are using a dermoscope. Computer aided diagnosis may also be applied on the digital images, and specialized optical methods may be used to see with higher resolution and at a greater penetration depth. In addition, the potential of DNA stripping methods and bioimpedance measurements have also been investigated. Reports suggest that ultrasound also has the potential to be an accurate diagnostic tool, with up to 99% sensitivity and specificity.

Recently, the use of ultrasound to study skin lesions has benefited from an increasingly concerted effort by physicians to standardize ultrasound examinations [5]. This is coupled with literature describing the differential appearance on ultrasound images of various skin lesions, chief among them melanoma, basal cell carcinoma (BCC), and benign nevi. Melanoma is generally described as being the most hypoechoic, with-well defined borders, a thin echo entry, and a spindle-like shape [6-9]. In contrast, BCC lesions are less hypoechoic, often feature many hyperechoic spots, and their contours are generally more irregular, while nevi are moderately hyperechoic, well demarcated from the dermis, and oval-to-circular in shape [6-10]. According to one of the co-authors K. Szalai (a radiologist with extensive experience in the ultrasound examination of skin lesions), the qualitative features of these lesions are as summarized in Table 1.

	Melanoma	BCC	Nevus
Shape	spindle	elongated/irregular	irregular
Attenuation (ref. dermis)	less	more	less
Homogeneous?	yes	no	no
Sharp, regular contour?	yes	no	no
Dendritic contour?	no	yes	no
Contains outer frame?	no	yes	no

Table 1. Qualitative features present in ultrasound images of melanoma, BCC, and nevus lesions in the experience of co-author K. Szalai.

As a means of comparing the ultrasound features quoted in the literature as well as those given in Table 1, Fig. 1 shows example ultrasound images of a melanoma, BCC, and nevus. In the melanoma lesion, vertical spreading of the hypoechoic, spindle-formed lesion with a thin echo entry is visible. The

BCC is seen as a lesion with a bright frame enveloping it, while the nevus has an irregular border and relatively heterogeneous inner lesion.
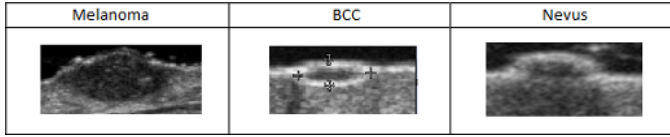

Figure 1.Representative ultrasound images of a melanoma, BCC, and nevus lesions. (For details on collection of data, see Methods section.)

The above features are arguably closely connected to the biophysical microstructure of the lesions, though these remain to be confirmed. For instance, several histologic features have been suggested for the common appearance of hyperechogenic spots in BCC lesions [10]. As other examples, the authors speculate that the spindle-like form of melanoma lesions on ultrasound images could be an indication of the nature of their lateral growth; and that the hyperechoic frame around the BCC lesions could be a sign of the palisading basal cells enveloping the lesion.

There are existing studies that investigate quantitative and semi-quantitative ultrasound markers to differentiate between common skin lesions, notably from the Institute of Cancer Research in London and the Institute of Fundamental Technological Research in Warsaw, in close collaboration with clinical partners. For some examples, see [11-13]. The current work is an attempt to investigate the relative usefulness of textural and shape markers, as inspired by clinical experience, in differentiating between melanoma, BCC, and benign nevi.

II.   METHODS

A. Data Collection

After obtaining approval from the local Ethics Committee, data were obtained from the Dermatooncology Department of Semmelweis University. As part of routine clinical practice, incoming patients suspected of having a cancerous skin lesion were examined using an ultrasound machine (HI VISION Preirus with 5-18 MHz EUP-L75 transducer, both from Hitachi, Tokyo, Japan) prior to surgical excision. In addition, healthy patients were also recruited in order to increase the number of nevi images.

Those patients who gave their informed consent had their images included in the current study, along with their final diagnosis. For those who had their suspicious lesion removed, the final diagnosis was made based on the histology. Images containing lesions 1-2 mm thick were selected (N=248), with 73 melanoma, 130 BCC, and 45 nevi.

Processing was carried out in MATLAB (Mathworks, Natick, MA). Prior to analysis, a MATLAB script removed personal data from the ultrasound image and a unique identification number linked the image with the final diagnosis.

B. Segmentation

After using the MATLAB command `imfill` to remove cross marks from the image, a semi-automated method was used to find the lesion, by manually clicking inside the lesion and then growing an active contour from the selected point to the inner edges of the lesion. The boundary layer of the lesion was then found by normalizing the brightness of the lesion center to 0 and the surrounding stroma to 1, and finding in a direction perpendicular to the lesion edge, the points that first reached 80% brightness. A post-lesion region and dermal region were defined immediately below and to the right of the lesion.

C. Features

The source images were B-mode images, which made the use of quantitative features difficult. Nevertheless, there was an effort to define the features in such a way as to make them as independent of the imaging platform as possible. To this end, the following 12 features were defined, with the first 5 being shape-based features, and the last 7 being texture-based features:

1. Lesion boundary circularity (LB-CIRC): 4π times the area of the lesion divided by the square of the perimeter. Multiplication by 4π ensures that the value is between 0 and 1, with a perfect circle having a value of 1. [14]

2. Lesion boundary curvature variability (LB-CUR-STD): first, for each point on the boundary, the point 3 points away on each side is taken, and the angles and distances to these points is considered. From this, the curvature of each point on the boundary is calculated [15], of which the standard deviation is taken.

3. Lesion boundary ellipse deviation variability (LB-ED-VAR): an ellipse is fitted to the lesion boundary, and the standard deviation of the distance between the fitted ellipse and the lesion boundary is calculated. [16]

4. Lesion boundary layer thickness (LBL-THICK-MU): the mean thickness of the boundary layer.

5. Lesion boundary layer thickness variability (LB-THICK-STD): the standard deviation of the thickness of the boundary layer.

6. Lesion echogenicity (L-ECHO-MU): the mean of the pixel values inside the lesion.

7. Lesion echo heterogeneity (L-ECHO-STD): the standard deviation of the pixel values inside the lesion.

8. Lesion boundary layer echogenicity (LBL-ECHO-MU): the mean pixel value of the boundary around the lesion.

9. Lesion boundary layer heterogeneity (LBL-ECHO-STD): the mean of the standard deviation of each boundary line layer emanating radially from the inner edge of the lesion.

10. Lesion boundary layer heterogeneity variability (LB-ECHO-STD2): the standard deviation of the standard deviation of each boundary layer line emanating radially from the inner edge of the lesion.

11. Post lesion echogenicity (pL-ECHO-MU): the mean pixel value of the post-lesion region relative to the dermal region.

12. Post lesion speckle size (pL-SS): the -6dB full width half-maximum of the mean transverse autocorrelation in the post-lesion region relative to the dermal region.

D. Classification Training and Testing

As a measure of diagnostic value – both of the features by themselves and of the classifiers – the area under the curve (AUC) measure was used, which refers to the area under the receiver operating characteristic (ROC) curve that plots sensitivity against 1-specificity [17]. The advantage of the AUC metric is that it does not require a threshold value to be chosen. For the classifiers, another metric was also employed, namely the specificity at 100% sensitivity, which refers to the percentage of negative diagnoses correctly made when all positive cases are found.

Two types of classifiers were trained: AdaBoost and Support Vector Machines (SVM). For cross-validation, 50 random equal populations of melanomas, BCC, nevi were generated, each of which were then 2-fold (holdout) cross-validated 50 times. The two aforementioned metrics were calculated for each of the resulting 2500 instances, investigating the ability of the classifier to differentiate between one type of lesion from the other two.

III. RESULTS AND DISCUSSION

Table 2 shows the highest attainable AUCs for each of the features. All features derived their highest AUCs from the "nevus versus other" classification, with the exception of pL-ECHO-MU, where it was attained for "BCC versus other". This suggests that most of the features considered in the current work are best suited to differentiating benign nevi from cancerous BCCs and melanoma. The most powerful features seem to be the LBL-ECHO-VAR, L-ECHO-VAR LB-ED-VAR features, which express in turn the heterogeneity of lesion and lesion boundary layer echogenicity, as well as the error of ellipsoid fitting to the lesion boundary. However, due to the small differences between AUC values, it is difficult to make far-reaching conclusions. Interestingly, the simplest (and perhaps most intuitive) feature, namely the mean lesion echogenicity, performs the worst.

Feature	Highest AUC
1. LB-CIRC	0.72
2. LB-CURV-VAR	0.76
3. LB-ED-VAR	0.78
4. LBL-THICK-MU	0.73
5. LBL-THICK-VAR	0.75
6. L-ECHO-MU	0.62
7. L-ECHO-VAR	0.79
8. LBL-ECHO-MU	0.73
9. LBL-ECHO-VAR	0.81
10. LBL-ECHO-VAR2	0.73
11. pL-ECHO-MU	0.77
12. pL-SS	0.74

Table 2. Highest attainable AUC for each feature.

Table 3 shows the classifier performance for all possible permutations of lesion groups. There is little difference between the performance of the SVM and AdaBoost. Although it is not possible to tell what the highest possible classifier performance could have been, the aforementioned small difference suggests that the values are close to the theoretical limits. The best performance is provided by a differentiation between BCCs and nevi, followed by the clinically much more useful differentiation between benign nevi and cancerous lesions (melanoma and BCC).

Although the high values of AUC suggest good differentiation, the value of specificity at 100% sensitivity provides a clinically more relevant question – for instance, in the case of nevus vs others, this value shows the number of correctly classified nevi while all cancerous lesions are detected. For both classifiers, this is about a fifth of all cases, which means that if the current method was used as a prescreening tool prior to observation by a dermatologist, about four-fifths of benign nevi would still be referred to the dermatologist if nearly all melanoma were to be spotted. At 91.1% sensitivity, 50.2% of nevi were spotted, which means that only half of the nevi would be referred to a specialist; however, this would come at the unacceptable cost of missing nearly 10% of melanoma lesions.

	SVM		AdaBoost	
	AUC	Specificity	AUC	Specificity
Nevus vs **others**	0.86	0.19	0.84	0.23
BCC vs **others**	0.80	0.15	0.81	0.20
MM vs **others**	0.59	0.04	0.62	0.03
BCC vs nevus	0.90	0.45	0.90	0.43
MM vs nevus	0.80	0.18	0.78	0.26
MM vs BCC	0.69	0.18	0.74	0.21

Table 2. Classifier performance for all possible permutations of lesion groups. Specificity is quoted for when sensitivity is set to 100% for the class in **bold**.

From the above results, it can be seen that although the results are promising, further work needs to establish the source of errors in the ultrasound classification to see if these can be remedied. One possible source of error is the segmentation algorithm, whose accuracy and robustness should be tested. Another way to improve diagnostic accuracy could be to take several images of the lesion, thereby providing more data points for the calculation of the features. Lastly, the sample size could be increased to provide for better classifier training as well as better evaluation of classifier performance.

IV. SUMMARY

The current work has shown that using a set of 12 ultrasound image features inspired by the observations of a radiologist, automated differentiation between nevi and cancerous lesions (melanoma, BCC) could be attained. An AUC of at least 0.84 for both classifiers was obtained, demonstrating the potential of ultrasound imaging for performing such differential diagnosis. However, at a specificity of at most 23% for 100% sensitivity, the current method is not yet clinically useful. Having more images of the same lesion, improving the segmentation algorithm, refining the feature definitions, as well as increasing the number of

training images, could all help in improving the method to the point of making it clinically useful.

REFERENCES

[1] G. P. Guy Jr., C. C. Thomas, T. Thompson, M. Watson, G. M. Massetti, and L. C. Richardson. "Vital signs: melanoma incidence and mortality trends and projections—United States, 1982–2030." MMWR Morb Mortal Wkly Rep 64, no. 21, pp. 591-596, 2015. http://www.cdc.gov/Mmwr/preview/mmwrhtml/mm6421a6.htm. Accessed Sept 4, 2016.

[2] Cancer Facts and Figures 2016. American Cancer Society. http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-047079.pdf. Accessed Sept 4, 2016.

[3] G.P. Guy Jr., S. Machlin, D. U. Ekwueme, K. R. Yabroff. "Prevalence and costs of skin cancer treatment in the US, 2002–2006 and 2007–2011." Am J Prev Med. vol. 48, pp. 183–7, 2015.

[4] D. S. Rigel, J. Russak, R. Friedman. "The evolution of melanoma diagnosis: 25 years beyond the ABCDs." CA Cancer J Clin vol. 60, pp. 301-316, 2010. http://onlinelibrary.wiley.com/doi/10.3322/caac.20074/full. Accessed Sept 4, 2016.

[5] X. Wortsman, F. Alfageme, G. Roustan, S. Arias-Santiago, A. Martorell, O. Catalano, M. S. di Santolo, K. Zarchi, M. Bouer, C. Gonzalez, R. Bard, "Guidelines for Performing Dermatologic Ultrasound Examinations by the DERMUS Group." Journal of Ultrasound in Medicine, vol. 35, no. 3, pp. 577-580, 2016.

[6] T. Cammarota, F. Pinto, A. Magliaro, A. Sarno. "Current uses of diagnostic high-frequency US in dermatology." Eur J Radiol, vol. 27, pp. S215-S223, 1998.

[7] L. Machet, F. Ossant, F. Patat, G. Georgesco, J. M. Grégoire, L. Vaillant, M. Samimi, M. Naouri, Y. Mourtada. High resolution ultrasound imaging of melanocytic and other pigmented lesions of the skin. INTECH Open Access Publisher, 2011.

[8] A. Mandava, P. R. Ravuri, R. Konathan. "High-resolution ultrasound imaging of cutaneous lesions." Indian J of Radiol Imaging vol. 23, no. 3, p. 269, 2013.

[9] X. Wortsman. "Sonography of the primary cutaneous melanoma: a review." Radiology research and practice, p. 814396, 2012.

[10] H. Uhara, K. Hayashi, H. Koga, T. Saida. "Multiple hypersonographic spots in basal cell carcinoma." Dermatologic Surgery vol. 33, no. 10, pp. 1215-1219, 2007.

[11] C. C. Harland, S. G. Kale, P. Jackson, P. S. Mortimer, J. C. Bamber. "Differentiation of common benign pigmented skin lesions from melanoma by high-resolution ultrasound." Brit J Derm, vol. 143, no. 2, pp. 281-289, 2000.

[12] H. Piotrzkowska-Wroblewska, J. Litniewski, E. Szymanska, A. Nowicki. "Quantitative Sonography of Basal Cell Carcinoma." Ult Med & Biol, vol. 41, no. 3, pp. 748-759, 2015.

[13] D. Rallan, N. L. Bush, J. C. Bamber, C. C. Harland. "Quantitative discrimination of pigmented lesions using three-dimensional high-resolution ultrasound reflex transmission imaging." J Invest Dermat vol. 127, no. 1, pp. 189-195, 2007.

[14] C. Di Ruberto, A. Dempster, "Circularity measures based on mathematical morphology." Electronics Letters, vol. 36, no. 20, p. 1, 2000.

[15] M. Worring and A. W. Smeulders, "Digital curvature estimation." CVGIP: Image Understanding, vol. 58, no. 3, pp. 366/382, 1993.

[16] Y-Y Liao, P-H Tsui, C-K Yeh. "Classification of benign and malignant breast tumors by ultrasound B-scan and Nakagami-based images." J. Med. Biol. Eng, vol. 30, no. 5, pp. 307-312, 2010.

[17] J. A. Hanley, B. J. McNeil. "The meaning and use of the area under a receiver operating characteristic (ROC) curve." Radiology, vol. 143, no. 1, pp. 29-36, 1982.