Dermatologist-like feature extraction from skin lesion for improved asymmetry classification in PH^2 Database

Rajib Chakravorty[1] and Sisi Liang[1] and Mani Abedini[1] and Rahil Garnavi[1]

***Abstract*— Asymmetry is one of key characteristics for early diagnosis of melanoma according to medical algorithms such as (ABCD, CASH etc.). Besides shape information, cues such as irregular distribution of colors and structures within the lesion area are assessed by dermatologists to determine lesion asymmetry. Motivated by the clinical practices, we have used Kullback-Leibler divergence of color histogram and Structural Similarity metric as a measures of these irregularities. We have presented performance of several classifiers using these features on publicly available PH^2 dataset. The obtained result shows better asymmetry classification than available literature. Besides being a new benchmark, the proposed technique can be used for early diagnosis of melanoma by both clinical experts and other automated diagnosis systems.**

***Index Terms*— Lesion Asymmetry, Melanoma, PH^2**

I. Introduction

Though melanoma is one of the most common and fatal type of skin cancer ([1], [2], [3]), early detection of melanoma can reduce mortality ([1]). Experts use clinical guidelines such as ABCD as proposed in [4], Menzies method [5], CASH algorithm [6], etc. to assess visual features in dermoscopic images for detecting melanoma lesions. In all of these approaches, several visual characteristics are detected according to established criterion. A numeric score is assigned to each of these characteristics and a final score is calculated to assess the risk of lesion being a melanoma. The ABCD rule guides on how to score a lesion based on its four visual features - 1) Asymmetry (A), 2) Border Irregularity (B), 3) Color (C) and 4) Diameter (D). The Total Dermoscopy Score (TDS) is given by weighted score of individual features (the weights are 1.3, 0.1, 0.5 and 0.5 respectively). Though the use of dermoscopy enhances sensitivity of disease detection, diagnosis of a skin lesion is not trivial even for health care professionals ([7], [8]). Automated analysis of lesion images removes the subjectivity from the process.

The automated classifiers available in literature are grouped in two broad categories - 1) Using the features as defined by medical algorithms (ABCD, CASH etc.) 2) Using the features directly from the image (texture, edges etc.). However, only the first approach is of practical value because the health care professional can use the medical algorithms to assist for the final decision making ([9]). Therefore, detection of medical characteristics in skin lesion is more important for early detection of melanoma.

[1]The authors are with the IBM Research, Australia, Level-5, 204 Lygon Street, Carlton, VIC 3053, {rachakra*,sisiliang,mabedini,rgarnavi@au1.ibm.com}

Asymmetry is considered to be a key characteristics of malignant skin lesion ([6], [5], [10]). This features is weighted maximally for calculating TDS score in ABCD rule ([4]). Besides shape asymmetry, dermatologists consider the asymmetry in color and structural distribution between two halves along each of the axis (mutually perpendicular) to determine the symmetry type of a lesion. In ABCD rule, assessment of these properties result in three classes as described below (example is given in figure 1):

- **Class 0 - Full Symmetry** - The distribution of shape, color and structural properties of the lesion is symmetric along both the axes
- **Class 1 - Half Symmetry** - The distribution of shape, color and structural properties of the lesion is symmetric along one of the axes and not the other
- **Class 2 - Full Asymmetry** - The distribution of shape, color and structural properties of the lesion is asymmetric along both the axes

Existing literature ([10], [11], [12], [13], [14], [15], [16]) processes the skin lesion image to extract various low level features that correspond to "asymmetry index". These indices or scores do not provide indication as to the class of symmetry as dictated by rules such as ABCD and rather are used as features for disease classification (i.e. melanoma or otherwise). Authors such as [17], [18] and [19] assess the asymmetry of the skin lesion, differently. The latter approach is particularly useful because it provides additional evidence to the nature of the lesion. This information can help a health care professional decide on the most suitable treatment plan ([20]).

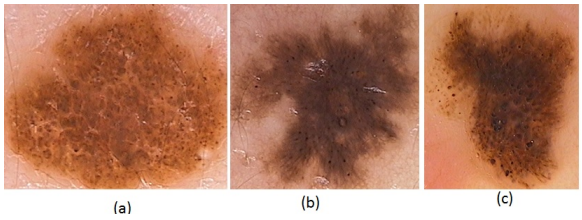

Fig. 1. Example of lesions of different symmetry types (a) Full Symmetry (b) Half Symmetry (c) Full Asymmetry

Several major drawbacks exist in the published literature in the domain of asymmetry assessment of skin lesions. These are:

- Various authors have used different data set to obtain results. This makes a systematic comparison of various algorithms difficult, if not impossible (This problem is

978-1-4577-0220-4/16/$31.00 ©2016 IEEE

observed by authors in d'Amico et. al, [10]).

- Where public dataset, such as PH^2, is used, some images were not analyzed. For example, a) the images containing noise (presence of hair etc.) ([19]) b) the images containing "Half symmetric" lesions ([18], [19])
- Most of the algorithms aim to solve the disease diagnosis problem and not asymmetry classification (Type 0, 1 and 2 in ABCD rule).

To this effect, we have proposed features that capture the asymmetric distribution of color and structure in the lesion. We used these features in several classifiers using proposed features. The proposed methodology is tested using the complete dataset of PH^2 ([21]). The obtained result overcomes the drawbacks described above and forms a new benchmark for asymmetry classification problem. Moreover, the improvement in asymmetry detection can contribute to better diagnosis of melanoma when used by both health care professionals or automated melanoma diagnosis classifiers.

The rest of the article is arranged as following: Section II explains the dataset and the steps of the algorithm. Section III describes the various implementation considerations and the performance of the proposed algorithm. Section IV highlights the findings and proposes some future directions.

II. Methodology

A. Dataset

In this article, we used the PH^2 database [21] that contains 200 dermoscopic images and includes medical annotation of all the images. Each image has a manually segmented lesion (in the form of a mask) and is categorized into one of the three clinically relevant symmetry types (0,1 and 2).

TABLE I

The distribution of the images in three categories of symmetry.

Symmetry Type	No. of images	Description
0	117	Fully Symmetric
1	31	Half symmetric along only one of the axes
2	52	Fully Asymmetric

We have used the expert segmented lesion to extract features and classification experiment. This avoids any additional sources of errors that may be introduced by an automated segmentation algorithm.

B. Symmetry type assessment

The workflow to assess the symmetry type is shown in figure 2.

1) Feature Extraction: The description of the features extracted is given in Table II and the process is briefly described in figure (3).

The extracted features are motivated by the workflow usually followed by dermatologist to assess symmetry types. The geometrical features (GF) captures the various shape related features of the lesion. However, more important

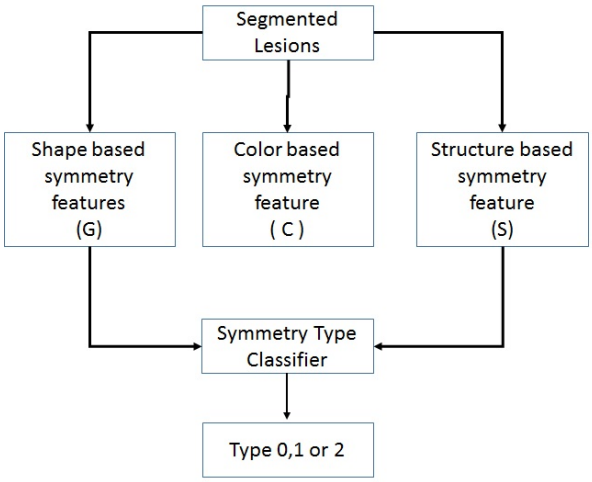

Fig. 2. The classification workflow: Each segmented lesion image is processed to capture the imbalance of shape, color and structure between each half along 2 axes of the lesion. The process gives rise to three vectors (G, C and S respectively) for each image. The vectors are combined, $F = G, C, S$ to be used as features for classification.

TABLE II

Feature extracted from images

Feature Type	Feature List	Description
Geometry Features (GF)	Solidity, Extent, Diameter, Circularity, Eccentricity, Aspect Ratio, Ratio of Major to Minor Axis	Extracted from the segmented mask image
Color Features (CF)	Kullback-Leibler divergence ([22]) of the color distribution (each channel separately) from the two halves along each axis.	Extracted from the segmented RGB image
Structural Features (SF)	Structural similarity ([23]) of two halves (with one half flipped along the axis of interest)	Extracted from the gray scale image of the segmented lesion

to the dermatologists (when assessing asymmetry) are the asymmetries of color and structures present in two halves along each of the axes. Kullback-Leibler (KL) measures the divergence ([22]) between two distributions. The histograms of each of the RGB channels from both of the halves are collected. KL divergence between two histograms provides a measure of color difference. Similarly, 'Structural symmetry' metric ([23]) considers the difference in luminance, contrast and structural composition of two images. Applying the methodology on each half of image indicates the structural difference needed for asymmetry assessment.

We implemented the feature extraction and classification in Python (`https://store.continuum.io/cshop/anaconda/`). Scikit-Image ([24]) package was used to extract the features (geometric and structural) from the dermoscopy images. Additionally, we used feature selection method to reduce the feature vector for each classifier to be used.

2) Classification: We tested a number of classifiers to detect the symmetry types of lesions. In order to make a valid comparison with the published result [17], we implemented

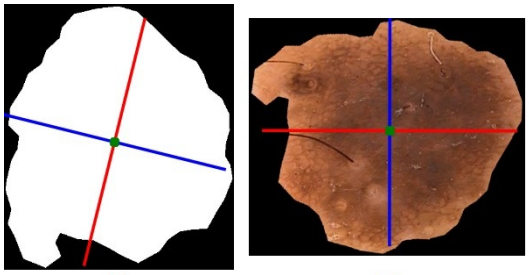

Fig. 3. (a) The major and minor axes obtained from the segmented mask image of the lesion. The shape asymmetry features are collected from this image. (b) The RGB lesion is rotated such that the major and minor axes align with the horizontal and vertical axes, respectively. The color and structural asymmetry features are collected from this image.

a binary classifier for each of the symmetry types (0, 1 and 2). The models were generated by classifiers in Leave-One-Out (LOO) fashion ([25]). The predicted results from each iteration of LOO was accumulated and the accuracy, AUC and precision were calculated.

Table I indicates that there is a severe imbalance of the dataset. The samples from symmetry type 0 is more than half of the entire dataset. Classifiers perform poorly in this scenario ([26]). We used SMOTE (`https://github.com/fmfn/UnbalancedDataset`) to generate synthetic dataset such that the number of positive and negative samples for each classifier become balanced.

III. Results and Discussions

A. Classification performance

The performance of the designed classifiers compared with that in [17]. The comparison is shown in tables III-V (the values in each cell is in X,Y format where X and Y denote the corresponding metric with and without feature selection regime respectively).

The performance of the proposed method can be directly compared with that from [17]. Asymmetry classification in [19] declares a slightly better performance (accuracy =0.87, sensitivity= 0.83, specificity=0.89, roc=0.86) where only a part of PH^2 data was used. However, the authors did not separate this partial result and therefore the proposed method could not be compared. Another asymmetry classification result in [18] declares an accuracy of 74% (the proposed method shows better performance in general) - however, the dataset used was different. [NA=Not Available]

B. Discussion

The proposed methodology outperforms the previously published result. Consistent classification performance in all the metrics across various classifiers indicates the suitability of the proposed features and methodologies. In general, the classifiers perform better with "feature selection" regime all of the metrics except 'Precision' where the best performance is shown by K-Nearest classifier without the feature selection for all three types of symmetries.

IV. Conclusion

Medical algorithms such as ABCD, CASH, Menzies method etc. provide important decision making input for health care professionals to diagnose melanoma at an early stage. Therefore, lesion properties such as "Asymmetry" need to be correctly classified by any assistive automated system such as Decision Support system. Motivated by dermatologist-like assessment of asymmetry, we have proposed a set of features to detect irregular distribution of color and structural properties of the lesion along with shape based features. Use of these features are demonstrated using the public dataset of PH^2 images. We tried a number of classifiers yielding improved asymmetry detection over published result on the same dataset. The images contain variances in brightness and lighting conditions and noises such as presence of hair etc. Therefore, the performance of the proposed methodology points to its robustness. The result presented in this article can be used as a new benchmark for the asymmetry classification problem. Moreover, incorporating such a system in melnaoma decision support workflow will be very valuable.

References

[1] A.C. Geller, S.M. Swetter, K. Brooks, M.F. Demierre, and A.L. Yaroch, "Screening, early detection and trends for melanoma: Current status (2000-2006) and future directions," *J. American Academy of Dermatology*, vol. 57, pp. 555–572, 2007.

[2] American Cancer Society, ," Available at `http://www.cancer.org/cancer/skincancer-melanoma/detailedguide/melanoma-skin-cancer-key-statistics`.

[3] Melanoma Institure Australia, ," Available at `http://www.melanoma.org.au/understanding-melanoma/melanoma-facts-and-statistics/`.

[4] S.W. Riemann, A.B. Cognetta, R. Talamini, R. Corona, and F. Sara et. al, "Abcd rule of dermatoscopy: a new method for early recognition of malignant melanoma," *Eur J. Dermatology*, vol. 4, pp. 521–527, 1994.

[5] S.W. Menzies, K.A. Crotty, C. Ingvar, and W.H. McCarthy, *Dermoscopy an Atlas*, 3rd edition edition.

[6] U. Weigert, W.H.C. Burgdorf, and Wilhelm Stolz, *ABCD Rule*, pp. 113–117, Second edition edition.

[7] I. Maglogiannis and C. Doukas, "Overview of advanced computer vision systems for skin lesions characterisation," *IEEE Transaction on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 721733, 2009.

[8] C.A. Morton and R.M. Mackie, "Clinical accuracy of the diagnosis of cutaneous malignant melanoma," *Br J Dermatol*, vol. 138, no. 2, pp. 283287, 1998.

[9] J.L.G. Arroyo and B. Zapirain, "Automated detection of melanoma in dermoscopic images," in *Computer Vision Techniques for the Diagnosis of Skin Cancer*, J. Scharcanski and M. E. Celebi, Eds., Series in BioEngineering, pp. 139–192. Springer Berlin Heidelberg, 2014.

[10] d'Amico M, M. Ferri, and I. Stanganelli, "Qualitative asymmetry measure for melanoma detection," in *Proc. IEEE Intl. Symposium on Biomedical Imaging*, 2004, vol. 2, pp. 1155–1158.

[11] W.V. Stoecker, W.W. Li, and R.H. Moss, "Automatic detection of asymmetry in skin tumours," *Computerized Medical Imaging and Graphics*, vol. 16, no. 3, pp. 191–197, 1992.

[12] V. Ng and D. Cheung, "Measuring asymmetries of skin lesions," in *Proc. IEEE International Conf. Computational Cybernetics and Simulation*, 1997, vol. 5, pp. 4211–4216.

[13] P. Schmid-Saugeon, "Symmetry axis computation for almost-symmetrical and asymmetrical objects : Application to pigmented skin lesions," *Medical Image Analysis*, vol. 4, pp. 269–282, 2000.

TABLE III

CLASSIFICATION PERFORMANCE OF FULL SYMMETRY

Classifier	Accuracy	Specificity	Sensitivity	Precision	AUC	F1 Score
Abedini et. el. [17]	0.66	NA	NA	0.69	0.65	NA
AdaBoost	0.75,0.71	0.67,0.62	**0.81**,0.78	0.74,0.68	**0.75**,0.71	**0.71**,0.67
SVM	0.74,0.72	0.66,0.63	0.80,0.80	0.73,0.68	0.74,0.72	0.70,0.69
Decision Tree	0.72,0.66	0.65,0.58	0.79,0.72	0.72,0.68	0.72,0.66	0.68,0.61
K-Nearest	0.72,0.71	0.64,0.67	0.78,0.74	0.72,**0.79**	0.71,0.69	0.67,0.63
Naive Bayes	**0.76**,0.72	**0.69**,0.67	0.80,0.74	0.77,0.78	**0.75**,0.71	**0.71**,0.65

TABLE IV

CLASSIFICATION PERFORMANCE OF HALF SYMMETRY

Classifier	Accuracy	Specificity	Sensitivity	Precision	AUC	F1 Score
Abedini et. el. [17]	**0.83**	NA	NA	0.14	0.49	NA
AdaBoost	**0.83**,0.78	0.89,0.89	**0.44**,0.34	0.39,0.45	0.65,0.65	**0.90**, 0.87
SVM	**0.83**,0.78	0.89,0.88	0.43,0.32	0.42,0.39	0.66,0.62	**0.90**,0.87
Decision Tree	0.81,0.75	0.90,0.88	0.39,0.28	0.45,0.39	0.66,0.60	0.88,0.85
K-Nearest	0.80,0.71	**0.91**,0.90	0.39,0.27	0.52,**0.55**	**0.68**,0.64	0.88,0.81
Naive Bayes	0.81,0.73	0.89,0.88	0.38,0.26	0.42,0.42	0.65,0.60	0.88,0.83

TABLE V

CLASSIFICATION PERFORMANCE OF FULL ASYMMETRY

Classifier	Accuracy	Specificity	Sensitivity	Precision	AUC	F1 Score
Abedini et. el. [17]	0.79	NA	NA	0.47	0.71	NA
AdaBoost	0.80,0.79	**0.90**,0.86	0.60,0.58	0.73,0.62	0.78,0.73	0.86,0.85
SVM	0.81,0.79	0.87,0.84	0.62,0.61	0.63,0.54	0.75,0.71	0.87,0.86
Decision Tree	**0.83**,0.72	0.89,0.85	**0.67**,0.46	0.69,0.62	**0.79**,0.68	**0.88**,0.80
K-Nearest	0.82,0.67	**0.90**,0.88	0.63,0.42	0.73,**0.75**	**0.79**,0.70	**0.88**,0.44
Naive Bayes	0.81,0.77	**0.90**,0.83	0.60,0.56	0.73,0.50	0.78,0.68	0.86,0.85

[14] S. Seidenari, G.Pellacani, and C. Grana, "Pigment distribution in melanocytic lesion images: a digital parameter to be employed for computer-aided diagnosis," *Skin research and Technology*, vol. 11, pp. 236–241, 2005.

[15] H. Iyatomi, H. Oka, M.E. Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka, and K.Ogawa, "An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm," *Computerized Medical Imaging and Graphics*, vol. 32, pp. 566–579, 2008.

[16] C. Barata, M. Ruela, M. Francisco, T. Mendonca, and J.S. Marques, "Two systems for the detection of melanomas in dermoscopy images using texture and color features," *IEEE Systems Journal*, vol. 8, no. 3, pp. 965–979, 2014.

[17] M. Abedini, Q. Chen, N.C.F. Codella, R. Garnavi, and X. Sun, "Accurate and scalable system for automatic detection of malignant melanoma," in *Dermoscopy Image Analysis*, M.E. Celebi, T. Mendonca, and J.S. Marques, Eds. CRC Press, 2014.

[18] M.J.M. Vasconcelos, L. Rosado, and M. Ferreira, "A new risk assessment methodology for dermoscopic skin lesion images," in *Proc. IEEE International Symposium on Medical Measurements and Applications*, Turin, Italy, 2015, pp. 570–575.

[19] M.J.M. Vasconcelos, L. Rosado, and M. Ferreira, "Principal axes-based asymmetry assessment methodology for skin lesion image analysis," in *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Ryan McMahan, Jason Jerald, Hui Zhang, StevenM. Drucker, Chandra Kambhamettu, Maha El Choubassi, Zhigang Deng, and Mark Carlson, Eds., vol. 8888 of *Lecture Notes in Computer Science*, pp. 21–31. Springer International Publishing, 2014.

[20] P. Schmid-Saugeon, J. Guillod, and J. Thiran, "Towards a computer-aided diagnosis system for pigmented skin lesions," *Computerized Medical Imaging and Graphics*, vol. 27, pp. 65–78, 2003.

[21] T. Mendona, P. M. Ferreira, J. Marques, A. R. S. Marcal, and J. Rozeira, "Ph^2 - a dermoscopic image database for research and benchmarking," in *Proc. IEEE International Conference in Medicine and Biology Society*, Osaka, Japan, 2013.

[22] S. Kullback and R.A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[23] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," vol. 13, no. 4, pp. 600–612, 2004.

[24] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors, "scikit-image: image processing in Python," *PeerJ*, vol. 2, pp. e453, 6 2014.

[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[26] H. He and E.A Garcia, "Learning from imbalanced data," vol. 21, no. 9, pp. 1263–1284, Sept. 2009.