Enhanced classification of malignant melanoma lesions via the integration of physiological features from dermatological photographs

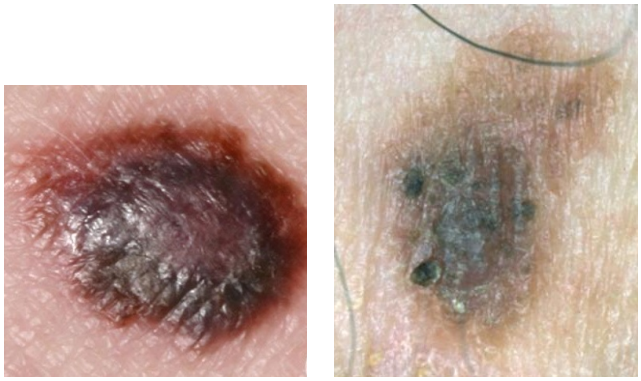
Shahid Haider, Daniel Cho, Robert Amelard, Alexander Wong, David A. Clausi

***Abstract*— Traditional methods for early detection of melanoma rely upon a dermatologist to visually assess a skin lesion using the ABCDE (Asymmetry, Border irregularity, Color variegation, Diameter, Evolution) criteria before confirmation can be done through biopsy by a pathologist. However, this visual assessment strategy taken by dermatologists is hampered by clinician subjectivity and suffers from low sensitivity. Computer-aided diagnostic methods based on dermatological photographs are being developed to aid in the melanoma diagnosis process, but most of these methods rely only on superficial, topographic features that can be limiting in characterizing melanoma. In this work, a hybrid feature model is introduced for characterizing skin lesions that combines low-level and high-level features, and augments them with a set of physiological features extracted from dermatological photographs using a nearest-neighbor nonlinear model to improve classification performance. The physiological features extracted from the lesion for the proposed hybrid feature model include those based on: i) eumelanin concentrations, ii) pheomelanin concentrations, and iii) blood oxygen saturation. The proposed hybrid feature model was evaluated on 206 dermatological photographs of skin lesions (119 confirmed melanoma cases, 87 confirmed non-melanoma cases) using a cross validation scheme. The experimental results show that the proposed hybrid feature model, with integrated physiological features, provided improved sensitivity, specificity, precision and accuracy for the purpose of melanoma classification.**

I. Introduction

Melanoma is the deadliest form of skin cancer [1] causing 65,000 global deaths in 2000 with 132,000 new cases occurring globally each year with the incidence rate predicted to only grow according to the World Health Organization [2]. If diagnosed early, simple extraction of the local cancerous tissue results in a high five-year survival rate [3]; if not, then the cancer can metastasize and prove fatal.

Dermatologists visually identify the pigmented lesion as cancerous using conventional dermatological photography, or photographs acquired using a specialized camera like a DermaScope. The method most commonly used to identify melanoma is for a dermatologist to assess them visually against the ABCDE criteria. The ABCDE criteria builds upon the ABCD criteria first introduced by Friendman *et al.* [4] in 1985 describing the observable features of melanoma by visual inspection, and was adopted quickly after introduction due to its simplicity. The mnemonic ABCD represents asymmetry (A), border irregularity (B), color variegation (C), and diameter (D) greater than 6 mm. Using the evolution (E) of the lesion over time as an observable indication of melanoma was incorporated into the criteria in 2004 [5] to form the final ABCDE criteria. Figure 1 shows example dermatological photographs of a lesion diagnosed with melanoma and a benign pigmented lesion. The ABCDE method, while being very simple, looks only at superficial, topological features and does not utilize more useful physiological biomarkers that characterize melanocytic lesions. It has also been known to suffer from clinician subjectivity leading to low sensitivity and specificity in skin cancer diagnosis [6].



(a) Melanoma (b) Non-Melanoma

Fig. 1: Examples of pigmented skin lesion images that have been diagnosed as (a) Melanoma and (b) Non-Melanoma.

In the hope of eliminating clinician subjectivity, computer-aided diagnosis methods are being developed to help improve diagnostic sensitivity. The key challenge though is to find the best set of features for characterizing melanoma. Two main categories of features have been proposed in previous literature: i) low-level features (LLFs), and high-level intuitive features (HLIFs). LLFs are simple image-related features that can be combined to provide a general characterization of skin lesions that are related to skin cancer diagnosis (e.g., asymmetry). For example, Cavalcanti and Scharcanski combined 11 low-level features (e.g., solidity, extent, circularity, equivalent diameter and axis length ratio) to describe the lesion's asymmetry [7]. These LLFs have the benefit of not requiring significant design time, since they are common application-agnostic features that are not specific to skin cancer analysis. However, low-level feature sets, due to their inherent high dimensionality, can suffer

*This work was supported by Agfa Healthcare, Ontario Centres of Excellence (OCE), National Natural Sciences and Engineering Research Council of Canada (NSERC), and Ontario Ministry of Economic Development and Innovation. This research was undertaken, in part, thanks to funding from the Canada Research Chairs program.

S. Haider, D. Cho, R. Amelard, A. Wong, and D. A. Clausi are with the Department of Systems Design Engineering, University of Waterloo, Ontario, Canada, N2L 3G1 {sa2haide, s34cho, ramelard,a28wong,dclausi}@uwaterloo.ca

from the curse of dimensionality, increased computational complexity, and possible over-fitting in sparse feature spaces during classification [8].

HLIFs are features that have been carefully designed and modelled to identify intuitive and semantically meaningful characteristics that a dermatologist can interpret (e.g., ABCDE). Amelard *et al.* were able to show a higher sensitivity in melanoma classification can be achieved via the use of HLIFs. Furthermore, HLIFs, when combined with LLFs, were shown to achieve a higher specificity and accuracy than LLFs alone [8]. On their own, the specificity and accuracy of the small feature set was low, however, it is shown that HLIFs can utilize less data to adequately populate a feature space and any suboptimal performance can be attributed to lack of data.

Similar to how validation and confident diagnosis is performed by pathologists by excising the lesion and checking for the non-superficial features such as the existence of cancerous cells and the level of spread, having feature sets that rely on superficial, topological features alone for computer-aided diagnosis of melanoma cannot be relied upon for high sensitivity diagnosis. Cavalcanti and Scharcanski have recognized this issue with existing feature models and have recently paired their LLFs with physiological features derived from dermatological photographs, which showed improved sensitivity to melanoma [9]. Inspired, by this approach of integrating physiological features with existing feature sets, the main contribution of this work is the creation of a hybrid feature model that combines LLFs and HLIFs, and integrates an extended set of physiological features that differ from those proposed by Cavalcanti and Scharcanski, to improve diagnostic sensitivity for the purpose of computer-aided melanoma diagnosis.

II. Hybrid Feature Model for Melanoma Analysis

The proposed hybrid feature model for enhanced classification of malignant cancerous lesions consists of three types of features: i) low-level features, ii) high-level features, and iii) physiological features. A description of each type of feature is provided below.

A. Low-Level Features

Using existing image processing techniques, numerous features can be extracted and combined together to provide an objective assessment of a skin lesion against the ABCD criteria. LLFs are those that aim to do just that. In the proposed hybrid model, the same feature extraction methods presented in Cavalcanti and Scharcanski [10] was used to extract 52 features from dermatological photographs:

- f_1: Solidity: the ratio between the lesion area and its convex hull area.
- f_2: Extent: the ratio between the lesion area and its bounding box area.
- f_3: Equivalent diameter.
- f_4: Circularity.
- f_5: The ratio between the principal axes.
- f_6: The ratio between sides of a bounding box containing the lesion.
- f_7: The ratio between the lesion perimeter and its area.
- f_8: The difference between the areas in the lesion that are divided by the major axis divided by the lesion area.
- f_9:The difference between the areas in the lesion that are divided by the minor axis divided by the lesion area.
- f_{10}: The ratio of the areas divided by the major axis.
- f_{11}:The ratio of the areas divided by the minor axis.
- f_{12-14}: The average gradient magnitude of the pixels in the dilated lesion rim, in each one of the three color channels.
- f_{15-17}: The variance of the gradient magnitude of the pixels in the dilated lesion rim, in each one of the three color channels.
- f_{18-20}: Dividing the lesion into 8 symmetric regions and computing the average gradient magnitudes across the dilated rim, in each of the three color channels.
- f_{21-23}: Dividing the lesion into 8 symmetric regions and computing variance of the gradient magnitudes across the dilated rim, in each of the three color channels.
- f_{24-27}: Maximum, minimum, mean and variance of the pixels intensities inside the lesion segment in the color variation channel.
- f_{28-39}: Maximum, minimum, mean and variance of the pixels intensities inside the lesion segment in each of the color channels.
- f_{40-42}: Ratios between mean values of the three original color channels.
- f_{43-48}: A count of the pixels who match the six hues typically associated with melanoma.
- f_{49-52}: The maximum, minimum, mean and variance of the pixels intensities inside the lesion segment to represent the textural variation.

B. High-Level Intuitive Features

Utilizing higher level features that have been specifically designed for the application of assessing a dermatological photograph of a lesion against the ABCD criteria has proven to improve classification sensitivity and also shown to populate its feature space efficiently [8]. In the proposed model, the four HLIFs with the most discriminating power presented by Amelard *et al.* [11] are used:

- f_1: Maximized color asymmetry score.
- f_2: Structural asymmetry score.
- f_3: The sum of the normalized differences of the lesion with a morphologically closed border and of the lesion with a morphologically opened border.
- f_4: The normalized absolute difference of the border and a low frequency representation of itself.

C. Physiological Features

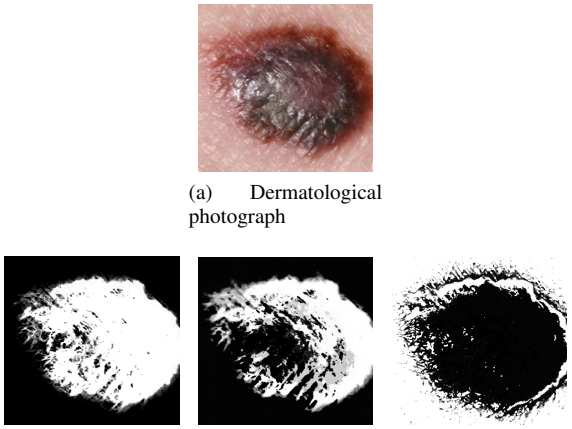
Incorporating physiological information as features can provide important clinical information. In order to design appropriate physiological features, we turn to how melanoma

forms. Cutaneous melanoma is characterized by the cancerous growth of melanin-producing cells called melanocytes. Melanocytes exist in the stratum basale of the epidermis. In cutaneous anatomy, "melanin" commonly refers to the subtypes eumelanin and pheomelanin, which contribute to an individual's skin color. Eumelanin is responsible for black-brown pigmentation, whereas pheomelanin is responsible for pink-red pigmentation. A higher overall melanin density results in darker skin, since melanin's molecular structure is particularly well-suited to absorbing ultraviolet and visible light [12]. Normally, the distribution of melanin is homogeneous over a small distance, resulting in uniform color. Melanoma, on the other hand, can result in a heterogeneous melanin density due to the local growth of melanosomes. In addition, due to the large collection of cells that exist in melanoma, angiogenesis occurs causing there to be a massive uptake of oxygenated blood to the area [13]. Inspired by this, we therefore turn to modeling not only eumelanin and pheomelanin concentrations, but also blood oxygen concentrations.

In the proposed hybrid model, we extend upon the nearest neighbor modeling strategy proposed in [9] to estimate the concentrations of eumelanin and pheomelanin, as well as the blood oxygen saturation inside the lesion on a pixel-by-pixel level (which is not done by [9]). Using a biophysical-based spectral model for simulating light interacting with human skin proposed by Krishnaswamy and Baranoski [14], we learn an extended nearest neighbor inverse model and use it to generate concentration maps for eumelanin, pheomelanin, and blood oxygen from dermatological photographs. Based on the generated concentration maps, the following physiological features were integrated into the proposed hybrid model:

- f_1: mean eumelanin concentration inside the lesion.
- f_2: mean pheomelanin concentration inside the lesion.
- f_3: variance of eumelanin concentration inside the lesion.
- f_4: variance of pheomelanin inside the lesion.
- f_5: mean blood oxygen saturation inside the lesion.

An example of the resultant concentration maps generated using the learned extended nearest neighbor inverse model is shown in Figure 2.



(a) Dermatological photograph

(b) Eumelanin map (c) Pheomelanin map (d) Blood oxygen sat. map

Fig. 2: The resultant concentration maps for eumelanin, pheomelanin, and blood oxygen saturation extracted from a dermatological photograph. Dark values represent a small concentration/saturation and brighter values represent a large concentration/saturation.

III. Results

A. Experimental Setup

We used a Bayesian classification scheme to assess the class separability of each feature set. In the Bayesian classification approach, each class (melanoma and non-melanoma) is modelled as a conditional multivariate normal distribution. This classification scheme was chosen to emphasize the robustness of the feature space rather than a particular classifier.

We constructed a dataset of 206 clinical images (119 confirmed melanoma cases, 87 confirmed non-melanoma cases) with dermatological photographs from DermIS [15] and DermQuest [16]. For each dermatological photograph, the following feature sets were extracted for forming the hybrid feature model:

- S_L: 52 LLFs [10].
- S_H: four HLIFs [11].
- S_P: five proposed physiological features.

We evaluated the following feature models (which comprise of different permutations of LLFs, HLIFs, and physiological features) for comparison:

- $S_L = S_L$
- $S_{LH} = S_L \bigcup S_H$
- $S_{LP} = S_L \bigcup S_P$
- $S_{LHP} = S_L \bigcup S_H \bigcup S_P$

We evaluated each feature set S using a cross-validation scheme. In particular, using the 206 images, 90% were randomly chosen for training and 10% for testing. This was repeated 50 times over each feature set with the arithmetic average of the metrics recorded.

B. Classification Results

The recorded classification results between the different feature sets are summarized in Table I. For each metric, the corresponding feature set that attains the highest score is boldface. S_{LH} outperforms S_L across all performance metrics, consistent with [11]. S_{LP} also outperforms S_L, indicating that incorporating physiological features adds discriminative information on top of LLFs. The best classification performance is attained by concatenating all the feature sets together (S_{LHP}) to form a hybrid feature model. This indicates that LLFs, HLIFs, and physiological features each add unique information to the feature extraction stage. This also indicates that the inclusion of physiological features in S_{LP} and S_{LHP} improves the classification performance over using just the topographic, superficial feature sets, S_L and S_{LH}.

TABLE I: Summary of classification results from the different feature models.

Feature set	Sensitivity		Specificity		Precision		Accuracy	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
S_L	84.95	10.97	70.69	14.61	80.93	9.18	79.00	6.90
S_{LP}	85.96	9.67	75.47	14.32	83.62	9.28	81.76	7.63
S_{LH}	85.94	10.43	72.07	13.31	81.45	9.35	80.10	6.71
S_{LHP}	**87.73**	9.30	**76.34**	13.66	**84.38**	9.12	**83.05**	7.11

C. Physiological Features Choice

Cavalcanti and Scharcanski chose to augment LLFs with physiological features representing the mean eumelanin and pheomelanin concentrations in the lesion [9], S_C. This greatly improved their specificity relative to the proposed LLFs, S_L from Table I, but hindered their specificity and precision. We chose to include the mean blood oxygen saturation in addition to the mean and variance of eumelanin and pheomelanin concentrations into S_P due to the angiogenesis present in melanoma development [13]. The result of including these additional physiological features to the LFFs is improved performance across all metrics. This is summarized in Table II with the highest scores in boldface. This demonstrates that the inclusion of additional physiological features introduced in the proposed hybrid feature model has the potential to improve melanoma classification over that proposed by Cavalcanti and Scharcanski [9].

TABLE II: Summary of classification results between Cavalcanti and Scharcanski [9](S_C) and S_{LP}.

Feature set	Sensitivity		Specificity		Precision		Accuracy	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
S_C	81.60	10.60	71.23	17.10	81.06	9.53	77.71	8.36
S_{LP}	**84.02**	8.33	**72.40**	16.58	**81.97**	10.11	**79.62**	7.82

IV. Conclusion

In this work, a hybrid feature model was presented for characterizing skin lesions that combines low-level and high-level features, and integrates a set of physiological features based on eumelanin concentrations, pehomelanin concentrations, and blood oxygen saturation extracted from dermatological photographs to improve classification performance. Experimental results show that the proposed feature model provided improved classification performance when compared to existing state-of-the-art methods. In the future, the proposed feature model can be extended to look at more complex physiological features that describe the normal homeostasis occurring in human skin, or the divergence from normal homeostasis that occurs in melanoma.

References

[1] A. F. Jerant, J. T. Johnson, C. Demastes Sheridan, and T. J. Caffrey, "Early detection and treatment of skin cancer." *American family physician*, vol. 62, no. 2, 2000.

[2] R. Lucas, T. McMichael, W. Smith, and B. Armstrong, *Solar ultraviolet radiation: global burden of disease from solar ultraviolet radiation*, ser. Environmental Burden of Disease Series, A. Pruss-Ustun, H. Zeeb, C. Mathers, and M. Repacholi, Eds. World Health Organization, 2006, no. 13.

[3] American Cancer Society, "Cancer Facts & Figures 2011," American Cancer Society, Atlanta, GA, Tech. Rep. ACSPC-029771, 2011.

[4] R. J. Friedman, D. S. Rigel, and A. W. Kopf, "Early detection of malignant melanoma: The role of physician examination and self-examination of the skin," *CA: a cancer journal for clinicians*, vol. 35, no. 3, pp. 130–151, 1985.

[5] N. R. Abbasi, H. M. Shaw, D. S. Rigel, R. J. Friedman, W. H. McCarthy, I. Osman, A. W. Kopf, and D. Polsky, "Early diagnosis of cutaneous melanoma: revisiting the abcd criteria," *JAMA: Journal of the American Medical Association*, vol. 292, no. 22, pp. 2771–2776, 2004.

[6] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara *et al.*, "Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet," *Journal of the American Academy of Dermatology*, vol. 48, no. 5, pp. 679–693, 2003.

[7] P. G. Cavalcanti and J. Scharcanski, "Automated prescreening of pigmented skin lesions using standard cameras," *Computerized Medical Imaging and Graphics*, vol. 35, no. 6, pp. 481–491, 2011.

[8] R. Amelard, A. Wong, and D. A. Clausi, "Extracting high-level intuitive features (hlif) for classifying skin lesions using standard camera images," in *Computer and Robot Vision (CRV), 2012 Ninth Conference on*. IEEE, 2012, pp. 396–403.

[9] P. G. Cavalcanti, J. Scharcanski, and G. V. Baranoski, "A two-stage approach for discriminating melanocytic skin lesions using standard cameras," *Expert Systems with Applications*, vol. 40, no. 10, pp. 4054–4064, 2013.

[10] P. G. Cavalcanti and J. Scharcanski, "Automated prescreening of pigmented skin lesions using standard cameras," *Computerized Medical Imaging and Graphics*, vol. 35, no. 6, pp. 481–491, 2011.

[11] R. Amelard, J. Glaister, A. Wong, and D. A. Clausi, "Melanoma decision support using lighting-corrected intuitive feature models," in *Computer Vision Techniques for the Diagnosis of Skin Cancer*, ser. Series in BioEngineering, J. Scharcanski and M. E. Celebi, Eds. Springer, 2013.

[12] R. Anderson and J. A. Parrish, "The optics of human skin." *Journal of Investigative Dermatology*, vol. 77, no. 1, 1981.

[13] J. Marcoval, A. Moreno, J. Graells, A. Vidal, J. M. Escribà, M. Garcia-Ramirez, and A. Fabra, "Angiogenesis and malignant melanoma. angiogenesis is related to the development of vertical (tumorigenic) growth phase," *Journal of Cutaneous Pathology*, vol. 24, no. 4, pp. 212–218, 1997.

[14] A. Krishnaswamy and G. V. Baranoski, "A biophysically-based spectral model of light interaction with human skin," in *Computer Graphics Forum*, vol. 23, no. 3. Wiley Online Library, 2004, pp. 331–340.

[15] Dermatology Information System, http://www.dermis.net, 2012, Accessed: 08 Nov 2012.

[16] DermQuest, http://www.dermquest.com, 2012, Accessed: 08 Nov 2012.