Computer Vision and Image Understanding 151 (2016) 124–137

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Clinically inspired analysis of dermoscopy images using a generative model

Catarina Barata [a,*], M. Emre Celebi [b], Jorge S. Marques [a], Jorge Rozeira [c]

[a] Institute for Systems and Robotics, Instituto Superio Técnico, Lisboa, Portugal
[b] Louisiana State University, Shreveport LA, USA
[c] Hospital Pedro Hispano, Matosinhos, Portugal

ARTICLE INFO

Article history:
Received 14 February 2015
Accepted 22 September 2015

Keywords:
Melanoma
Dermoscopy
Computer-aided diagnosis
Color detection
ABCD rule
Correspondence-LDA
Von-Mises-Gaussian distribution

ABSTRACT

Dermatologists often prefer clinically oriented Computer Aided Diagnosis (CAD) Systems that provide medical justifications for the estimated diagnosis. The development of such systems is hampered by the lack of detailed image annotations (medical labels and segmentations of the associated regions). In most cases, we only have access to weakly annotated images (text labels) that are not sufficient to learn proper models. In this work we address this issue and propose a CAD System that uses medically inspired color information to diagnose skin lesions. We deal with the weakly annotated dermoscopy images using the Correspondence-LDA algorithm to learn a probabilistic model. The algorithm is applied with success to the identification of relevant colors in dermoscopy images, obtaining an average Precision of 83.8% and a Recall of 89.8%. The proposed color representation is then used to classify skin lesions, resulting in a Sensitivity of 77.6% and Specificity of 73.0% using Random Forests, and a Sensitivity of 75.1% and Specificity of 77.5% using SVM. These results comparable favorably with related works.

1. Introduction

The most dangerous characteristic of melanoma is its ability to grow and spread rapidly to other tissues and organs [1]. This makes melanoma the deadliest form of skin cancer, although it is by far one of the less common types of skin related neoplasms. According to the most recent data, the incidence rate of melanoma has been steadily increasing for the past three decades and it currently ranks in the ninth position among the most common types of cancer in Europe alone [2]. The advanced stage of melanoma is often incurable and leads to the death of the patient, but an early diagnosis of this disease (when the abnormal growing cells are still contained within skin tissue) can lead to a full recovery [1]. Thus, a great effort has been put on the development of skin lesion visualization and diagnosis techniques, that can help dermatologists improve their diagnostic accuracies.

Dermoscopy is among the most popular imaging methods used by dermatologists, because it combines magnification and special illumination techniques that render an improved image of the skin lesion [3]. With this method, dermatologists are able to observe and analyze surface and subsurface structures that are invisible to the naked eye [1,4]. The observed structures, called dermoscopic criteria, play an important role in the diagnosis of melanoma and are considered in different medical procedures, such as the ABCD rule [5] and 7-point checklist [6]. The main drawback of dermoscopy is that it can only be effectively applied by trained practitioners [7]. Other negative characteristics of this method are its subjectivity and lack of reproducibility [8]. These drawbacks fostered the development of Computer Aided Diagnosis (CAD) systems, such as the ones described in [9–12] (see [13] for a survey on this topic), that can act as a second opinion tool and be used by non-experienced dermatologists [14].

Despite the interesting experimental results achieved by some of the CAD systems, dermatologists have pointed out that several of them have not been designed to work as a support tool [15]. The practitioners see these systems more as parallel/second opinion tools that give an output of melanoma or benign, without providing comprehensive medical information to justify the diagnosis. This black box structure and lack of interaction are two of the main reasons why dermatologists avoid including CAD systems in their routine practices. These two issues can be addressed with the development of more clinically oriented systems that focus not only on the diagnosis but also on the identification of key dermoscopic criteria (e.g., clinically relevant colors).

The development of clinically oriented CAD systems is an active topic of research. Different research groups have proposed strategies to detect the presence of dermoscopic criteria, such as pigment network [16–20], streaks [21,22], dots [23], and colors [24–27] or color

* Corresponding author.
E-mail address: ana.c.fidalgo.barata@ist.utl.pt (C. Barata).

http://dx.doi.org/10.1016/j.cviu.2015.09.011
1077-3142/

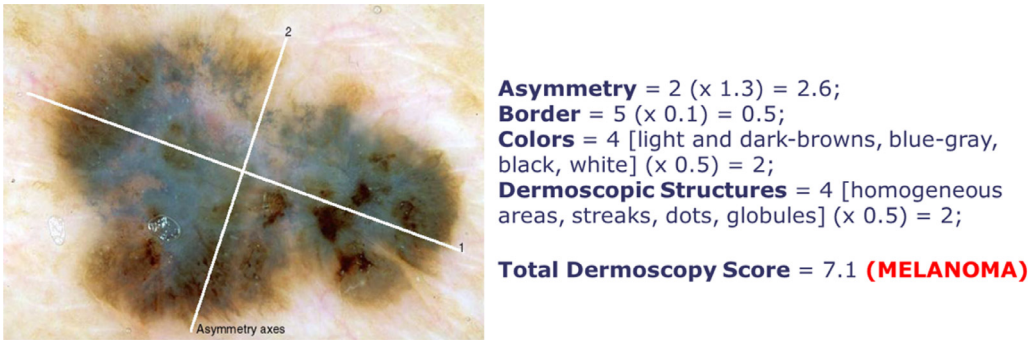

Fig. 1. Example of the application of the ABCD rule [1].

related structures [26–31]. However, only a few of these works use the detected criteria to obtain a diagnosis of the lesions as melanoma or benign [13], which demonstrates the difficulty of this problem.

One of the major challenges of developing a clinically oriented CAD system is that it might require a large number of detailed annotated images (i.e., medical text labels and segmentations of the relevant dermoscopic criteria). This detailed medical information must be obtained through consultation with experienced practitioners. Dermatologists usually provide text labels stating whether a dermoscopic criterion is present or absent, but do not perform their corresponding segmentations because it is a time consuming and subjective task. However, several of the methods described in the literature require detailed annotations (e.g., detection of colors [24–26], blue-whitish veil [28,30], and global patterns [32–34]), and can result in incomplete systems if the number of available segmentations is not sufficient. This limitation can be addressed through the design of systems that are capable of dealing with weakly annotated data (i.e., images for which there are text labels and it is not known which are the image regions that correspond to those labels). Such systems must be able not only to reproduce the medical labeling process in new images but also to identify the regions within the lesions that correspond to the text labels. Although one might argue that this last aspect is unnecessary as the system already provides text labels, it can be quite useful for dermatologists as it would allow them to associate the text outputs of the system with specific areas in the lesion and verify if the suggested output makes sense. It is also important for the designed systems to be able to diagnose the lesions as melanoma or benign using the detected medical criteria.

This work addresses the aforementioned problems and investigates the development of a clinically oriented CAD system, in which it is possible to learn a probabilistic model to represent the dermoscopic criteria using only medical text labels. The system is capable of i) reproducing the labeling process; ii) identifying the regions in the lesion associated with each of the labels, and iii) diagnosing the lesion as melanoma or benign. Various dermoscopic criteria could be used to study the labeling process. In this work we have selected the clinically relevant colors that are considered in the ABCD rule (Dark and Light Browns, Blue-Gray, Black, Red, and White) [5]. The selection of the color criterion is based both on the difficulty of the problem and on the fact that color detection systems usually require training examples of color segmentations. The probabilistic model used to learn the correlation between medical labels and image regions is Correspondence-LDA (corr-LDA) [35]. To the best of our knowledge this is the first time that such a model and approach are applied to the analysis and classification of dermoscopy images.

The paper is organized as follows. First we give an overview of the problem and the notation used (Section 2). Then, we discuss the state-of-the-art in annotation (Section 3), describe the probabilistic model (Section 4) and present the proposed modifications (Section 5). We discuss different possibilities to diagnose the skin lesions using the detected color information in Section 6. Finally we present the obtained results (Section 8) and conclude the paper (Section 9).

2. Problem formulation

2.1. Clinical analysis

A clinically oriented CAD system for the diagnosis of melanoma must have the following framework: i) identify relevant regions in the dermoscopy images and associate them with the dermoscopic criteria; ii) provide labels for the entire image stating whether the dermoscopic criteria are present or absent; and iii) use the identified medical information to estimate a diagnosis.

The first challenge that we must address is the selection of the dermoscopic criteria that must be identified by the developed CAD system. Medical procedures such as the ABCD rule [5] provide us with the necessary information regarding which are the criteria that dermatologists use to distinguish between benign lesions and melanomas. ABCD rule is a scoring approach that considers four different aspects of the lesion in order to obtain a diagnosis. The assessed criteria are: (A)symmetry regarding shape, color, and structures; irregular (B)orders; the number of (C)olors (up to six); and the existence of (D)ermoscopic structures, such as pigment network or streaks. During the diagnosis, dermatologists start by assigning an individual score to each of these criteria. Then, the scores are combined into a total lesion score using a weighted sum. The obtained score gives information about the level of suspiciousness/malignancy of the lesion. Fig. 1 shows an example of the ABCD rule [1].

In this work we address the detection of the clinically relevant colors considered by in the ABCD rule: Dark and Light Browns, Blue-Gray, Black, Red, and White (Fig. 2 shows some examples of lesions and the colors identified by experts). The detection of colors in dermoscopy images has already been addressed by some research groups [24–27,36]. Among these works, some require training examples of segmented color regions, which are not easy to obtain as was pointed out in the beginning of this paper. Other works do not use training examples and focus on the process of color quantification [27,36], usually using clustering methods, without actually identifying which are the colors that can be found in a given lesion. Our objective is to perform color identification and quantification first and then use this information to diagnose skin lesions. The main limitation is the lack of training examples of segmented color regions, since the segmentation of colors in dermoscopy images is a cumbersome and subjective task that is avoided by most dermatologists. Thus, we must investigate an alternative strategy that allows us to train a color model based on the available data.

2.2. Preliminary information and goals

Our dataset comprises D dermoscopy images in which the lesions were divided into small non-overlapping square patches, as shown