Expert Systems With Applications 56 (2016) 300–309

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Segmentation of melanocytic skin lesions using feature learning and dictionaries☆

Eliezer Flores, Jacob Scharcanski*

Instituto de Informática, Universidade Federal do Rio Grande do Sul., Avenida Bento Gonçalves 9500, RS 91501-970, Brazil

ARTICLE INFO

Article history:
Received 25 December 2014
Revised 29 January 2016
Accepted 21 February 2016
Available online 2 March 2016

Keywords:
Segmentation
Skin lesion
Standard camera
Dictionary learning
Feature learning
Sparse representation

ABSTRACT

Pre-screening systems for the diagnosis of melanocytic skin lesions depend of the proper segmentation of the image region affected by the lesion. This paper proposes a feature learning scheme that finds relevant features for skin lesion image segmentation. This work introduces a new unsupervised dictionary learning method, namely Unsupervised Information-Theoretic Dictionary Learning (UITDL), and discusses how it can be applied in the segmentation of skin lesions in macroscopic images. The UITDL approach is adaptive and tends to be robust to outliers in the training data, and consists of two main stages. In the first stage, a textural variation image is used to construct an initial feature dictionary and an initial sparse representation via Non-Negative Matrix Factorization (NMF). In the second stage, the feature dictionary is optimized by selecting adaptively the number of dictionary atoms. The greedy approach used for dictionary optimization is quite efficient and flexible enough to be applied to other dictionary learning problems. Furthermore, the proposed method can be easily extended for other image segmentation problems. The experimental results suggest that the proposed approach potentially can provide more accurate skin lesion segmentation results than comparable state-of-the-art methods. The proposed segmentation method could help to improve the performance of pre-screening systems for melanocytic skin lesions, which can affect positively the quality of the early diagnosis provided to skin lesion patients.

1. Introduction

Malignant melanocytic skin lesion is a type of skin cancer that attacks the melanocytes (the cells that produce the pigmented melanin, which colors skin, hair and eyes) (Melanoma Research Foundation, 2014). Although it is less common than other types of skin cancer, more than 100, 000 cases occur around the world every year, and this malignant skin lesion is very dangerous since it causes nearly 77% of skin cancer deaths (University of California, San Francisco, 2014). This type of skin cancer often is curable if detected in its early stages, but it is more likely than the other skin cancers to spread to other parts of the body (and potentially becoming lethal) if not treated early (Skin Cancer Foundation, 2014). A malignant melanocytic skin lesion can be classified into one of the following stages: “Stage I”, “Stage II”, “Stage III” and “Stage IV”. Here, the “Stage I” corresponds to the earliest detection stage and the “Stage IV” corresponds to the latest one. If detected still in “Stage I”, the 5-years survival rate is around 97%. However, the 5-years survival rate decreases to around 15% if the detection occurs when the skin lesion is in “Stage IV” (American Cancer Society, 2014).

Digital dermatoscope is the tool usually used to detect malignant melanocytic skin lesions. However, even among dermatologists, it has been observed that the use of a digital dermatoscope is not yet a common practice. A cross-sectional survey (Engasser & Warshaw, 2010) of all US fellows of the American Academy of Dermatology shows that only 48% of respondents use a dermatoscope for skin lesion diagnosis. Furthermore, in some countries, scheduling an appointment with a dermatologist may require long waiting times. Therefore, there is interest in developing pre-screening systems to identify malignant melanocytic skin lesions using standard camera images (i.e., simple photographs, without special lighting or dedicated equipment). These systems are not designed to replace the dermatologist, but can be helpful in tasks such as pre-screening or establishing a priority order for specialized consultations (Massone, Wurm, Hofmann-Wellenhof, & Soyer, 2008; Wong, Scharcanski, & Fieguth, 2011).

Skin lesion pre-screening can be organized in four steps (Cavalcanti, Scharcanski, & Baranoski, 2013): (a) pre-processing,

☆ The authors thank CNPq – Brazilian National Council of Scientific and Technological Development – for funding this project.

* Corresponding author. Tel.: +55 5133087128.
E-mail addresses: esflores@inf.ufrgs.br (E. Flores), jacobs@inf.ufrgs.br (J. Scharcanski).

http://dx.doi.org/10.1016/j.eswa.2016.02.044
0957-4174/

where the input image is processed to facilitate the skin lesion segmentation; (b) segmentation, where the goal is to delimit the skin lesion region in the input image; (c) feature extraction, where the segmented skin lesion is represented by its features; and (d) classification, where the skin lesion is classified as benign or malignant based on the lesion features extracted previously. This paper deals with steps (a) and (b), with a primary focus on step (b), i.e., skin lesion segmentation using standard camera images.

Segmentation is an important step for extracting information from images and videos (Scharcanski, Shen, & Silva, 1993; Silva & Scharcanski, 2010; Wong et al., 2011), where the goal is to group similar pixels into regions, which is essential for performing several tasks in medical image analysis (Jung & Scharcanski, 2002; Siqueira, Scharcanski, & Navaux, 2002; Welfer, Scharcanski, & Marinho, 2011). Clustering based methods are possibly the most popular approaches for image segmentation (Oliver, Munoz, Batlle, Pacheco, & Freixenet, 2006; Scharcanski et al., 1993). However, the high dimensionality of modern image datasets poses a considerable challenge for clustering approaches (Boutsidis, Drineas, & Mahoney, 2009). Thus, a feature learning stage may be used to identify the relevant features that shall be used in the classification stage.

Among the most popular methods for unsupervised feature learning are PCA (Jolliffe, 2002), ICA (Hyvärinen, Karhunen, & Oja, 2001) and NMF (Lee & Seung, 2001). Given a set of input features, these methods learn a set of basis vectors and represent the input features as linear combinations of these basis vectors. Usually, NMF provides a sparse representation for the feature set (unlike PCA and ICA), since NMF allows only non-negative contributions, and there is a relationship between non-negativity and sparsity (explained using the active-set theory in optimization, or by a Bernoulli prior in Bayesian inference) (Li & Ngom, 2013). Sparse representations provide a parsimonious representations of the data, which is a fundamental principle of modern feature learning and classification (Bach, Jenatton, Mairal, & Obozinski, 2012).

A summary of the state-of-the-art in computer vision techniques for the diagnosis of skin cancer can be found in (Scharcanski & Celebi, 2013), and several algorithms have been proposed for skin lesion segmentation on standard camera images in the literature (Cudek, Grzymala-Busse, & Hippe, 2010; Manousaki et al., 2006; Ruiz, Berenguer, Soriano, & Martin, 2008; Tabatabaie, Esteki, & Toossi, 2009). Cavalcanti, Yari, and Scharcanski (2010b) proposed to segment skin lesions using the ***R*** channel of the ***RGB*** color space. Alcón et al. (2009) determined the non-lesion pixels by assuming a Gaussian-like distribution in the grayscale histogram. Tang (2009) proposed an elaborate segmentation scheme based on Anisotropic Diffusion (Perona & Malik, 1990) and a modified version of Gradient Vector Flow Snake (Xu & Prince, 1998). Cavalcanti, Scharcanski, Di Persia, and Milone (2011) proposed to use ICA to provide an initial segmentation, and then use the Chan–Vese method (Chan, Sandberg, & Vese, 2000) to obtain the final segmentation. Cavalcanti and Scharcanski (2011) proposed to use Otsu's Thresholding (Otsu, 1979) in a suitable multi-channel representation, which tries to maximize the discrimination between lesion and non-lesion skin regions. This multi-channel representation is also used by Cavalcanti, Scharcanski, Martinez, and Di Persia (2014), where new features are learned by NMF and the segmentation is performed by k-means (Kanungo et al., 2002). It has been proposed in Flores and Scharcanski (2014) to select atoms from a NMF-learned dictionary using an unsupervised version of the Information-Theoretic Dictionary Learning method (Qiu, Patel, & Chellappa, 2014). According to this method, the final skin lesion segmentation is obtained using the Normalized Graph Cut (NGC) method (Shi & Malik, 2000). However, we propose in our work a novel dictionary learning scheme that is more robust to outliers. The proposed dictionary learning scheme uses a textural variation image to initialize the dictionary, which is more efficient to initialize the dictionary than the multi-channel representation used in Flores and Scharcanski (2014) (see Section 4). Afterwards, the learned dictionary is optimized by selecting adaptively the number of dictionary atoms. In our proposed scheme, the number of dictionary atoms selected usually differs for different images, unlike in Flores and Scharcanski (2014) that uses a constant number of dictionary atoms for any skin lesion image (i.e., the number of dictionary atoms is an input parameter of that method). Based on our experiments, the proposed segmentation scheme tends to outperform comparable state-of-the-art skin lesion segmentation methods in terms of the XOR error criterion (including Flores & Scharcanski, 2014, see Section 4).

The thresholding based methods (Cavalcanti & Scharcanski, 2011; Ruiz et al., 2008) often rely on global skin lesion image statistical characteristics to discriminate the different image regions, but image regions with similar intensities are easily confused by thresholding based methods. On the other hand, active contour based methods (Cavalcanti et al., 2011; Tang, 2009) tend to overcome this common limitation of thresholding based methods, but are computationally more expensive and often tend to include nearby artifacts in the segmented skin lesion (e.g., hair). Clustering based methods (Cavalcanti et al., 2014; Flores & Scharcanski, 2014) often rely on global image features to segment skin lesions, and tend to discriminate similar adjacent regions while discarding nearby artifacts at significant computational cost. The proposed skin lesion segmentation method tends to be computationally more efficient than the active contour based methods, but less efficient than the thresholding based methods, and relies on the following assumptions: (1) the main skin lesion segment is the focus of attention, and is positioned centrally in the image; (2) the skin lesion does not touch any of the input image corners (to avoid interfering with the shading attenuation process); and (3) the majority of the image pixels are non-lesion (i.e., healthy skin). If these assumptions are not satisfied, the performances of pre-processing and post-processing stages could be negatively affected, leading to a sub-optimal overall system performance (see Section 4). In practice, these image acquisition constraints can be easily satisfied by taking simple precautions during the image acquisition step.

The remaining of this paper is organized as follows. Section 2 presents the proposed feature learning scheme. Section 3 shows how to segment the macroscopic images with the learned features. Section 4 discusses our experimental results, and the conclusions are presented in Section 5.

2. Proposed feature learning scheme for segmenting pigmented melanocytic skin lesions

Standard camera images of pigmented melanocytic skin lesions usually contain shading areas (artifacts) that may be confused with skin lesions. In medical image analysis, a pre-processing stage is often used to remove noise and artifacts (Jung & Scharcanski, 2004). Therefore, to avoid that such artifacts mislead the segmentation process, in our approach the input image is first shading attenuated by a pre-processing step (Cavalcanti, Scharcanski, & Lopes, 2010a), as detailed in Section 2.1. The proposed feature learning scheme starts by mapping the input image to a feature space that helps maximize the discrimination between lesion and non-lesion skin regions. Texture is generally regarded as a powerful feature for image analysis and segmentation (Cavalcanti & Scharcanski, 2011; Scharcanski, 2005; Scharcanski & Dodson, 2000). It has been shown that a textural variation image can be obtained for an input skin lesion image (Cavalcanti & Scharcanski, 2011), which provides discriminative features for skin lesion segmentation (see Section 2.2). Also, NMF has been successfully used for skin lesion