SCALE INVARIANT DESCRIPTORS IN PATTERN ANALYSIS OF MELANOCYTIC LESIONS

Carlos S. Mendoza, Carmen Serrano, Begoña Acha

Universidad de Sevilla, Escuela Superior de Ingenieros
Camino de los Descubrimientos s/n, 41092 Sevilla, Spain

ABSTRACT

In this paper we introduce the importance of scale invariance in properly discriminating some of the typical patterns found in melanocytic lesions, by dermatoscopic image analysis. Pattern discrimination is a necessary step before pattern irregularity (an indicator of malignancy) can be quantified. We propose a set of features that allows for the discrimination of such patterns even when they appear in different degrees of magnification. We show how an automated feature selection stage produces a preferred scale invariant set of features among non-invariant features, yielding the best classification rate for those features. The average correct classification rate for the five kinds of classified patterns rises up to 94%.

Index Terms— Biomedical image processing, Pattern classification, Skin, Medical diagnosis.

1. INTRODUCTION

In the last two decades a rising incidence of malignant melanoma has been observed. Because of a lack of adequate therapies for metastatic melanoma, the best treatment is still early diagnosis and prompt surgical excision of the primary cancer. Dermoscopy (also known as epiluminescence microscopy) is an in vivo method that has been reported to be a useful tool for the early recognition of malignant melanoma [1]. Its use increases diagnostic accuracy between 5 and 30% in clinical visual inspection [2].

Currently available digital dermoscopic systems offer the possibility of computer storage and retrieval of dermoscopic images. Some systems even display the potential for Computer Assisted Diagnosis (CAD) [3, 4]. As diagnostic accuracy with dermoscopy has been shown to depend on the experience of the dermatologist, CAD systems will help less-experienced dermatologists and provide a lower impact for inter-subject variability.

Most technical papers developing methods to automatically classify dermatologic images are based on the ABCD rule (Asymmetry, Border irregularity, Color variation, Diameter greater than 6mm or growing). Frequently, such papers present an approach to cover one or several letters of the rule, that is, some are based on detecting Asymmetry [5, 6], Borders [7–10], Color [11, 12] or Diameter [12].

One of the key ideas of this paper is that it is not focused on detecting specific features in the images to cover the four letters of the ABCD rule, but instead follows the new tendency in Dermatology: to look for specific patterns in the lesions which can lead physicians to an assessment, by means of the so-called pattern turbulence analysis. Before any pattern turbulence or irregularity can be analyzed, a previous requirement is lesion segmentation into single-pattern regions. Sometimes this can be done manually [13], but ultimately a fully automatic method must rely on proper pattern segmentation. The different patterns that can be found inside a lesion are: reticelar; globular; cobblestone; homogeneous; starburst; and parallel [14].

This pattern-oriented approach has been adopted by several authors [14, 15]. In previous work [14], Serrano and Acha modeled melanocytic texture \( L^*a^*b^* \) color planes using Markov Random Fields (MRF). Independently, Tanaka et al. [15] studied a considerable number of classical statistical measures for texture discrimination and implemented several feature selection techniques, using only the luminance channel.

One of the key issues that previous work has not addressed is that of scale invariance of descriptors, and its impact on appropriate generalization to natural pattern variations in clinical melanocytic samples. Under the assumption that scale invariance could account for differences in lesion magnification under non-standard acquisition procedures, we wanted to show whether specially designed scale invariant descriptors could perform better than previous ones.

In Serrano et al. the proposed model turns out to be rather good for discriminating 5 out of the 6 previously enumerated pattern classes. However, the MRF approach is difficult to extend to scale invariance. Despite their interesting ideas and adequate experimental results, Tanaka et al. do not provide information on which features get actually selected, preventing from further extensions towards a pattern segmentation method. However, their technique lends itself naturally to the study of scale invariance and its influence on classification success.

In this paper we propose a set of scale invariant features...
for melanocytic lesion discrimination that throws very high classification rates. In Section 2, we will revisit some of the classical measures suggested in the work by Tanaka et al. [15]. More precisely, our study relies on the shape of the connected components of pattern samples. We will produce a set of measures for the connected components representation of every pattern sample. Some of them will be scale invariant and some of them will not. In Section 3 we will show how the scale invariant measures in the connected components sample patterns are best suited for pattern discrimination, among the many measures proposed. Very good classification results for the selected descriptors account for the suitability of the scale invariance property. Finally, in Section 4 we will provide some conclusions and future lines of work on the issue.

2. CONNECTED COMPONENTS PATTERN CLASSIFICATION

The goal of the reported method is to create a set of descriptors suitable for skin pattern discrimination. In this paper we present a supervised method of classification, with a previous training stage. We are going to classify the following patterns: reticular, globular, cobblestone, homogeneous and parallel. A training set with images representing each pattern individually is available for training and classification. It is not the purpose of this paper to separate the lesion from the normal skin. Instead the input to the algorithm will be a manually segmented sample of a given manually classified pattern. Therefore, our database is formed by 40×40 color images, taken uniformly from all of the considered pattern classes.

2.1. Connected Components Extraction

In order to capture the nature of the different patterns, we will quantify the shape of the periodically-repeating pigmented structures that can be seen in the samples. To achieve that quantification a previous step is necessary, in which the pattern sample has to be transformed into a binary image. Then, on the binary image, connected (black and white) component analysis can be performed.

After transforming the sample images from their original RGB color space into a better, perceptually uniform $L^*a^*b^*$, we extract the luminance $L^*$ channel for further analysis. We will proceed smoothing the sample using a very small (3×3) averaging mask in order to avoid some of the noise present in the image and then threshold the resulting smooth image. We will use local thresholding, with different threshold values for changing positions in the sample. More precisely, we will use a further smoothed version of the sample (using a moving average filter of size 20×20) to define the threshold image. The size of this filter was determined experimentally, by visual inspection we confirmed that this value can vary slightly without significant changes in the result of the thresholding. Values in the sample that show higher intensity than the threshold image in the corresponding position will be set to 1 in the resulting binary image. See Fig. 1 for an illustrative example.

2.2. Connected Components Morphological Processing

Once a binary representation of the sample texture has been obtained, and prior to shape quantification we introduce a novel step in order to provide better behavior for the features that are to be computed. Since our classification is going to be based on statistical analysis of the shape of the connected components, we wish to isolate single blob-like components rather than having only one main region. In that way, many components will be available in each sample, and the statistical moments of shape can be computed robustly enough.

In order to guarantee a good number of blobs we introduce a preprocessing step based on morphological bridging of the binary samples [16]. Morphological bridging consists of eliminating the separation between non-connected components, creating bridges between previously connected pixels, that is, setting 0-valued pixels to 1 if they have two nonzero 8-neighbors that are not 8-connected, and thus separating connected 0-valued regions.

Since we are going to be working with both the white and the black connected component images for out pattern classification features, we proceed by applying morphological bridging to the binary samples and their complemented (logi-
2.3. Feature Extraction and Classification

Once we have computed the black and white blob images, feature extraction can take place. For each connected component in each one of the images we calculate its area and perimeter in pixels, and also two quantities called area-perimeter quotient and roundness quotient [15]. The analytical expression for the last measure is indicated in the following eq.:

\[
\text{roundness}(i) = \frac{4\pi \times \text{area}(i)}{\text{perimeter}(i)}
\]

(1)

Then, we compute the median, mean and variance for the white components and the black components, of the aforementioned four measures. That results in a 24 feature vector for each pattern sample.

Since our goal is to determine which features best discriminate our five possible classes, we wanted to evaluate only the relevance of feature choice. For this initial approach a feature-space nearest neighbor classification method has been used.

2.4. Feature Selection

Performing classification on all possible feature combinations using 24 features is untractable. One possibility is using the well known and preferred technique of forward/backward sequential feature selection. Forward feature selection consists of adding features one by one, in such a way that we will always add the greatest classification rate increasing features out of the ones that have not been added yet. We stop the process when no improvement is perceived or no features are left. Backward feature selection is analogous, we start classifying with the whole set of features and start removing one by one, in such a way that we always remove the feature whose removal translates in the best possible classification rate. We stop when no improvement is possible or no features are left to be removed. Comparing the results of both selections, we can keep the best combination of the two.

One of our main goals in this paper is to decide whether scale invariance is a desirable property of a set of features for abnormal skin pattern discrimination. The previously described feature selection step will help us find the answer.

3. EXPERIMENTAL RESULTS

The proposed algorithm has been tested on a database containing 100 40×40 image samples of the following patterns: reticular, globular, cobblestone, homogeneous and parallel. For each type of pattern 20 images have been used. These images were provided by the Dermatology Unit at Virgen del Roco Hospital (Sevilla, Spain). All of them were taken with a dermoscope Fotofinder, Schuco International London Limited. We have performed 10-fold cross-validation [17]: 90% of the total set of images has been used to train (90 images) and 10% to validate (10 images). The total set has been divided into 10 groups for the testing. Each time one different testing group has been employed to validate the algorithm. In this way, 18 images of each pattern have been used to train and two images of each type have been used to validate.

This cross-validation process has been repeated for each feature selection step. The best classification rate was obtained for the combination of the following features: mean of the white roundness quotient, mean of the black roundness quotient, variance of the white roundness quotient, variance of the black roundness quotient. All the mentioned features are somewhat scale invariant as compared to other unselected ones like area or perimeter. The classification average success rate was 94%. See Table 1 for more details.

The program has been implemented in Matlab. The computational time for the 10-fold cross-validation is 2.1s.

4. CONCLUSIONS AND FUTURE WORK

In this paper a method for melanocytic pattern discrimination has been presented. Other studies on the issue did not ap-
proach the issue of scale invariance. Our method, not only discriminates patterns with great accuracy, but also suggests the preference of scale invariant parameters over scale dependent ones.

In the future, better scale invariant features should be developed in order to separate melanocytic lesion into its elemental parts. Malignancy assessment based on pattern analysis requires pattern irregularity measures. These measures should account for the number of available patterns, and also for turbulences in the scale of patterns. For this analysis fully automatic pattern segmentation methods need to be developed. By all means, such segmentation techniques must be able to identify different scales of the same pattern before pattern turbulence can be analyzed.

In future work also the segmentation of the lesion from healthy skin must be obtained, as part of the pattern segmentation process. Thus, healthy skin patterns, and also the starburst pattern need to be discriminated. The starburst pattern was not analyzed in this paper because its description implies segmentation in itself, as its main discriminative characteristic from the rest of patterns is located in the edge of the starburst lesion.

Further on, we plan to develop a series of subjective experiments in order to properly convey the perceptual sensations that dermatologists evaluate when clinically assessing malignant melanocytic lesions. This is a necessary step towards automatic diagnosis, and is by no means a resolved topic.

5. REFERENCES


