Applied Soft Computing 64 (2018) 341–355

Contents lists available at ScienceDirect

Applied Soft Computing

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

Partial order label decomposition approaches for melanoma diagnosis

Javier Sánchez-Monedero [a,*], María Pérez-Ortiz [a], Aurora Sáez [b], Pedro Antonio Gutiérrez [c], César Hervás-Martínez [c]

[a] Department of Quantitative Methods, Universidad Loyola Andalucía, 14004 Córdoba, Spain
[b] Signal Theory and Communications Department, University of Seville, 41092 Seville, Spain
[c] Department of Computer Science and Numerical Analysis, University of Córdoba, 14071 Córdoba, Spain

ARTICLE INFO

Article history:
Received 19 December 2016
Received in revised form 25 October 2017
Accepted 27 November 2017
Available online 14 December 2017

Keywords:
Melanoma
Computer vision
Machine learning
Ordinal classification
Partial order
Skin cancer

ABSTRACT

Melanoma is a type of cancer that develops from the pigment-containing cells known as melanocytes. Usually occurring on the skin, early detection and diagnosis is strongly related to survival rates. Melanoma recognition is a challenging task that nowadays is performed by well trained dermatologists who may produce varying diagnosis due to the task complexity. This motivates the development of automated diagnosis tools, in spite of the inherent difficulties (intra-class variation, visual similarity between melanoma and non-melanoma lesions, among others). In the present work, we propose a system combining image analysis and machine learning to detect melanoma presence and severity. The severity is assessed in terms of melanoma thickness, which is measured by the Breslow index. Previous works mainly focus on the binary problem of detecting the presence of the melanoma. However, the system proposed in this paper goes a step further by also considering the stage of the lesion in the classification task. To do so, we extract 100 features that consider the shape, colour, pigment network and texture of the benign and malignant lesions. The problem is tackled as a five-class classification problem, where the first class represents benign lesions, and the remaining four classes represent the different stages of the melanoma (via the Breslow index). Based on the problem definition, we identify the learning setting as a partial order problem, in which the patterns belonging to the different melanoma stages present an order relationship, but where there is no order arrangement with respect to the benign lesions. Under this assumption about the class topology, we design several proposals to exploit this structure and improve data preprocessing. In this sense, we experimentally demonstrate that those proposals exploiting the partial order assumption achieve better performance than 12 baseline nominal and ordinal classifiers (including a deep learning model) which do not consider this partial order. To deal with class imbalance, we additionally propose specific over-sampling techniques that consider the structure of the problem for the creation of synthetic patterns. The experimental study is carried out with clinician-curated images from the Interactive Atlas of Dermoscopy, which eases reproducibility of experiments. Concerning the results obtained, in spite of having augmented the complexity of the classification problem with more classes, the performance of our proposals in the binary problem is similar to the one reported in the literature.

1. Introduction

Melanoma is a type of cancer that arises from the pigment-containing cells known as melanocytes. The most common type, the cutaneous melanoma, occurs on the skin. In Europe, approximately 100,000 cases are yearly diagnosed, with a death ratio around 13% [1]. Patient prognosis depends directly on tumour thickness, where mortality can be reduced to a great extent by early detection and diagnosis [2].

To improve survival rates, melanoma must be detected before the tumour has penetrated the epidermis (i.e. before the thickness is higher than 0.76 mm). In the case of early detection, the five-year survival rate is about 99%, otherwise dropping to 15% for patients with advanced disease [3]. The current detection process consists on a visual inspection by trained professionals using a dermatoscope, and the prognosis is evaluated measuring the depth of the melanoma by means of a biopsy. Dermatologists perform this manual visual inspection from dermoscopy images, but this process is time-consuming and error-prone, and it can lead to widely vary-

* Corresponding author at: Department of Quantitative Methods, Universidad Loyola Andalucía, Calle Escritor Castilla Aguayo, 4, 14004 Córdoba, Spain.
E-mail address: jsanchez@uloyola.es (J. Sánchez-Monedero).

https://doi.org/10.1016/j.asoc.2017.11.042
1568-4946/

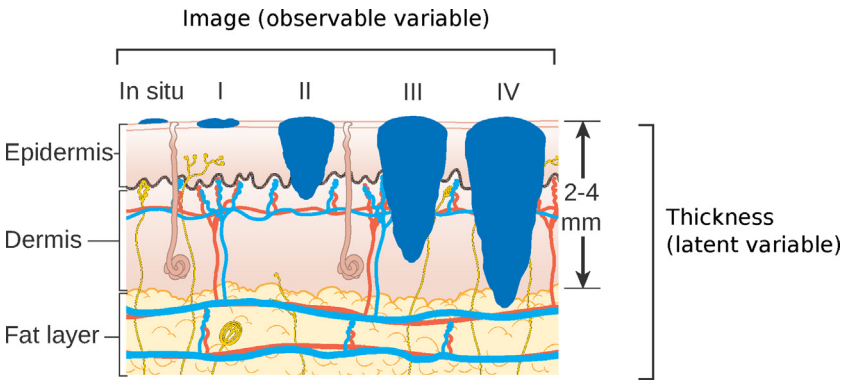

Fig. 1. Graphical representation of the different stages of melanoma, where both the observable data (dermoscopic image) and the unobservable or latent variable (thickness of the tumour) can be analysed.
Image credit: Cancer Research UK/Wikimedia Commons.

ing diagnosis. This motivates automated diagnosed methods [4,5]. Recent works propose new tools to aid or to improve this process [3], mainly based on dermoscopic image analysis. Although there are different lines of undergoing research (e.g. those based on skin temperature variations in the lesion), image analysis methods present the advantage of being cheaper and relatively easy to combine with existing detection procedures.

In the last years, computerised dermoscopy image analysis systems have been proposed to assist pigmented lesions diagnosis [6]. The majority of these works focus on the distinction of melanomas from benign lesions [7–9]. However, a finer grain classification is required for appropriate prognosis. The scarcity of studies on this topic and its inherent difficulty makes it a promising line for research. The first work in that line is the characterisation of two types of melanoma based on their thickness [10]. This study uses 49 features related to colour, geometry and texture, extracted from a private database of 141 images obtained with a company proprietary hardware system. Moreover, a recent study [11] focuses on the classification of three degrees of thickness for melanomas, but it excludes their distinction from benign lesions, which is crucial for constructing a complete detection tool.

In this paper, we propose to simultaneously address the problem of melanoma detection and thickness estimation within a five-class classification problem. To do so, we combine image analysis and machine learning procedures. Now, we summarise the feature extraction process and describe the dataset characteristics which motivate the development of specific machine learning methods. Particularly, the challenging issues found in this problem are: (1) the structure and topology of the classes and (2) the imbalanced nature of the classes that can bias classification performance in favour of majority classes.

Concerning the image analysis, we propose a set of 100 input features to describe images. The extracted features correspond to visual characteristics based on dermatologists clinical findings (see Section 4). Melanoma cases are distinguished from non-melanoma ones using the ABCD method, based on four clinical characteristics that describe a malignant melanoma: asymmetry (A), border irregularity (B), colour variegation (C) and differential structures (D). The rest of the features selected are related to melanoma thickness estimation, and analogously they are based on clinical criteria with respect to visual characteristics present in dermoscopic images [12,11].

When attempting to estimate the severity of a melanoma, it can be seen that the classes are imbued with order information. The Breslow index is modelled as an unobservable latent variable that represents the thickness of the tumour using the dermoscopic image (independent variable). Such latent variable can only be directly observed when performing a biopsy, in which case the actual tumour thickness can be measured and used to validate the prediction. Since the different Breslow index levels correspond to thresholds of the thickness, the corresponding class labels show an order relationship, in such a way that stage II melanomas are thicker than stage I ones, stage III implies a thicker lesion than stage II, and so on. Fig. 1 shows the different stages of a melanoma and analyses the observed and latent variable concepts in the frame of this problem. Please note that in this work we group stages III and IV due to the fact that they have similar clinical properties. This type of problems are known as ordinal classification problems, also referred to as ordinal regression [13]. They differ from nominal (standard) classification problems in the fact that there is an order arrangement between the categories, and they are different from regression because the distance between the values of the dependent variable (the class) is generally unknown. The most common situation in ordinal regression is that the categories come from the discretisation of a latent variable [13], which is exactly the case of the different stages of melanoma. Ordinal methods exploit the ordered nature of the classes to improve learners at the same time that penalise the magnitude of the classification errors (for example, in our case, misclassifying a stage 0 melanoma with a stage I should not be considered the same than confusing it with a stage III melanoma). Ordinal classification has been successfully applied to different areas such as Alzheimer' progression estimation [14] or sovereign ratings [15], among others. Section 2 provides some basic background on ordinal classification.

However, although this order is clear for the different stages of the melanoma (since they reflect different levels of thickness), it cannot be assumed for the benign lesion class. In this sense, the problem can be considered as a partially ordered classification task, for which we propose several machine learning strategies to exploit this characteristic. Fig. 2 illustrates the concept of partial order in a two-dimensional dataset, where it can be seen that $\mathcal{C}_1$ does not follow an order with respect to the rest of classes, while the rest are ordered in the input space ($\mathcal{C}_2$ is closer to $\mathcal{C}_3$ than to $\mathcal{C}_4$, and so on). Note that this structure can be found in very different classification problems, e.g. in medicine (non-disease vs. disease grades). In this case, it can be seen that a unique linear projection (which takes the order of the classes into account) is not feasible, while two projections (one for tackling the binary problem and other for the ordinal one) could separate the data satisfactorily. Ordinal problems with specific data structures (e.g. partial order problems or circular ordinal regression [16,17]) or other more complex label structures (such as multiple output ordinal regression, graded multilabel classification [18] or label ranking [19]) are recently receiving attention from the machine learning community.