

Selection of Image Features based on a Wrapper Model for the Classification of Skin Lesions

Roberta Barbosa Oliveira, João Manuel R. S. Tavares
Instituto de Engenharia Mecânica e Gestão Industrial, Faculdade de Engenharia,
Universidade do Porto, Rua Dr. Roberto Frias, 4200-465, Porto, Portugal,
emails: {pro12003, tavares}@fe.up.pt

Aledir Silveira Pereira
Departamento de Ciências de Computação e Estatística,
Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista
Rua Cristóvão Colombo, 2265, 15054-000, São José do Rio Preto, Brasil
email: aledir@ibilce.unesp.br

The number of skin cancer cases has been increasing every year. Thus, the development of automated computational systems to assist dermatologists becomes more and more imperative. In fact, the use of these systems can prevent more efficiently the development of skin cancer, as the dermatologists can use such systems to identifying relevant features of the lesions in the images under analysis. In the systems that have been developed, the computational classifiers have been predominant.

An attractive classifier for skin lesions from images must be efficient and fast. A relevant problem that affects the performance of the classifiers concerns the selection of features that are representative of the problem involved. A solution for this problem can be the use of a feature selection model, i.e. to use a data mining technique to discard the features that are redundant and reduce the classifier performance.

There are many models proposed for feature selection, such as filtering, wrapper, embedded and hybrid models. The wrapper model allows to select the most relevant features for classification, regardless of the chosen classifier learning algorithm. Basically, this model involves a search algorithm, an evaluation function and a performance function, which evaluates the results of the search algorithm and of the evaluation function. The search algorithm allows to select the best features for a given classification problem. After, each features subset established by the search algorithm is evaluated by the learning algorithm used in order to find the most relevant subset. Many classifiers can be used to evaluate the selected features and conduct the training of the classifier from training samples that must be classified into the appropriate classes.

As such, in this work, the combination of some search and learning algorithms based on the wrapper model for the features selection was studied. The model used

allowed the selection of relevant features from the training set and the discarding of the irrelevant features for the classification problem. Thus, it was attained an improvement in the computational classification of skin lesions.

Acknowledgements

The first author would like to thank “Conselho Nacional de Desenvolvimento Científico e Tecnológico”, in Brazil, for her PhD grant. This work was partially done in the scope of the research project PTDC/EEA-CRO/103320/2008, financially supported by “Fundação para a Ciência e a Tecnologia”, in Portugal.

References:

- I. Guyon; A. Elisseeff. An Introduction to variable and feature selection. *Journal of Machine Learning Research* 3:1157-1182, 2003.
- S. F. da Silva; A. Traina; M. X. Ribeiro; J. do E. S. Batista Neto; C. Traina. Ranking evaluation functions to improve genetic feature selection in content-based image retrieval of mammograms. *IEEE International Symposium on Computer-Based Medical Systems*, pp. 1-8, 2009.