11th World Congress on Computational Mechanics (WCCM XI)
5th European Conference on Computational Mechanics (ECCM V)
6th European Conference on Computational Fluid Dynamics (ECFD VI)
July 20 - 25, 2014, Barcelona, Spain

FEATURES SELECTION FOR THE CLASSIFICATION OF SKIN LESIONS FROM IMAGES

Roberta Barbosa Oliveira[1], Aledir Silveira Pereira[2] and João Manuel R. S. Tavares[1]

[1] Instituto de Engenharia Mecânica e Gestão Industrial,
Faculdade de Engenharia, Universidade do Porto,
Rua Dr. Roberto Frias, 4200-465, Porto, Portugal,
emails: {roberta.oliveira, tavares}@fe.up.pt

[2] 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

Key Words: *Features Selection, Classification, Filter Model, Skin Lesions.*

The interest in the development of medical image analysis systems to assist the medical diagnosis has been gained importance in the recent years. Particularly, their use by dermatologists has been supporting better and earlier diagnosis and consequently, improving the prevention to skin cancer, which has a very high incidence in the modern society.

One of the problems in the classification of skin lesions in images is the selection of the best set of representative features. In fact, many possible features can be irrelevant for the successfully diagnosis, and the complexity of the classification problem is proportional to the number of used features, so it is important to select the features that are really relevant to this problem of image-based diagnosis.

Here, we applied different methods based on the filter model to evaluate which is the best set of features to be used in the classification of skin lesions in images. The methods used were Correlation-based Feature Selection (CFS), Gain Ratio and ReliefF. The evaluated set was defined by a respective search algorithm, mainly Genetic Algorithms and Ranker. The filter model permits to evaluate the features without the need to use learning algorithms. Each set was generated and evaluated by the filter based on a statistic analysis. Then, the set was compared with the best set previously found. If the new set was considered the best one, it becomes the best set. This process was run in cycle until an established stopping criterion was reached.

The skin features used were: asymmetry, border, colour, distances, signature and statistic measures, which were computed from the original images. In the classification process, different predictors, such as, k-Nearest Neighbour (K-NN), Naïve Bayes and J48, were used and compared. The skin image database used is composed by 377 images, and the best performance was achieved by CFS evaluator filter, using genetic algorithms in the searching for the best features set, and K-NN classifier to classify the testing lesions as benign or malign.

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 with reference PTDC/BBB-BMD/3088/2012, financially supported by "Fundação para a Ciência e a Tecnologia", in Portugal.

REFERENCES

[1] R. C. Gonzalez and R. E. Woods. *Digital image processing*. 2. ed. New Jersey: Prentice Hall, 2002. 793 p.

[2] I. Guyon and A. Elisseeff. An Introduction to variable and feature selection. *Journal of Machine Learning Research 3*, p. 1157-1182, 2003.

[3] H. Liu and L. Yu. Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Engineering*, v. 17, n. 4, p. 491-502, 2005.