Combining Different Normalizations in Lesion Diagnostics

M. Skurichina[1], A.I. Belousov[2], D.C.G. de Veld[3,4], R.P.W. Duin[1], M.J.H. Witjes[3], H.J.C.M. Sterenborg[4], J.L.N. Roodenburg[3]

[1]Pattern Recognition Group, Department of Applied Physics, Faculty of Applied Sciences, Delft University of Technology, P.O. Box 5046, 2600GA Delft, The Netherlands
marina@ph.tn.tudelft.nl

[2]University of Muenster, Institute of Anorganical and Analytical Chemistry, Muenster, Germany

[3]Department of Oral and Maxillofacial Surgery, University Hospital of Groningen, Groningen, The Netherlands

[4]Photodynamic Therapy and Optical Spectroscopy Programme, Department of Radiation Oncology, Erasmus Medical Center - Daniel den Hoed Cancer Center, Rotterdam, The Netherlands

Abstract

In lesion diagnostics, it is sometimes hard to choose which data normalization is the best among the other ones. They all give a similar performance, but the information retrieved from the data can be different. In such a case, a combined decision of several classifiers constructed on differently normalized data can be used. Our simulation study performed on the autofluorescence spectra measured in the oral cavity shows that combining different data normalization techniques is useful, when a trained combiner is used as an aggregating rule for combining the classifiers.

1. Introduction

In the diagnosis of potentially cancerous lesions, one is interested in an accurate faultless detection of lesions. This can be done only when an accurate representation of healthy and diseased tissues is provided. However, in practice when taking the measurements, large variations within one group (class) of objects (e.g. healthy tissues or diseased tissues) may occur which are caused by the measuring equipment and/or the diversity of measured objects. To get rid of unnecessary deviations in the data description it might be useful to normalize the data, making representations of objects belonging to the same data class similar to each other. Usually when normalizing the data, some general information about the data classes is retrieved. But some part of the useful information can be lost. Different normalization techniques retrieve different information about the data classes. Sometimes it is quite hard to single out a normalization technique that ensures the best performance for a particular dataset. In such a case, it might be beneficial to combine several normalization techniques when solving the problem. One way to do this is to use a combined decision of several classifiers constructed on differently normalized data instead of a single decision of the classifier obtained on the uniquely normalized data.

In this paper we test this idea for lesion diagnostics performed on the autofluorescence spectra measured in the oral cavity [1]. We study the performance of single classifiers constructed on differently normalized data and the combined decision of these classifiers. This study is carried out for a 2-class problem. We consider non-normalized autofluorescence spectra and three types of spectra normalization: Unit Area normalization [2], Standard Normal Variate transformation [3,4] and Savitzky-Golay smoothing and differentiation [4,5]. To evaluate the quality of lesion classification, the Regularized Linear Classifier [6] and the Regularized Quadratic Classifier [6] assuming normal class densities are used. We also investigate the effectiveness of different combining rules when aggregating the classifiers.

2. Data and Normalization Techniques

The data consist of the autofluorescence spectra acquired from healthy and diseased mucosa in the oral cavity. The measurements were performed at the Department of Oral and Maxillofacial Surgery of the University Hospital of Groningen [1]. Autofluorescence spectra were collected from 97 volunteers with no clinically observable lesions of the oral mucosa and 137 patients having lesions in oral cavity. The measurements were taken at 11 different anatomical locations with excitation wavelength 365 nm. After preprocessing [1] each spectrum consists of 199 bins (pixels/wavelengths). In total, 857 spectra representing healthy tissue and 112 spectra representing diseased tissue were obtained.

For spectra normalization three normalization techniques are used:

• Unit Area normalization (UA)

$$a_i^{UA} = \frac{a_i}{U}, \quad U = \sum_{j=1}^{199} a_j, \quad i = \overline{1, 199},$$

where a_i is an intensity of a spectrum $\boldsymbol{A}=\{a_1, ..., a_{199}\}$ at bin i, i=1,...,199.

• Standard Normal Variate transformation (SNV) [3,4]

$$a_i^{SNV} = \frac{a_i - M}{S}, \quad i = \overline{1, 199},$$

where M is the mean and S is the standard deviation of a spectrum $\boldsymbol{A}$.

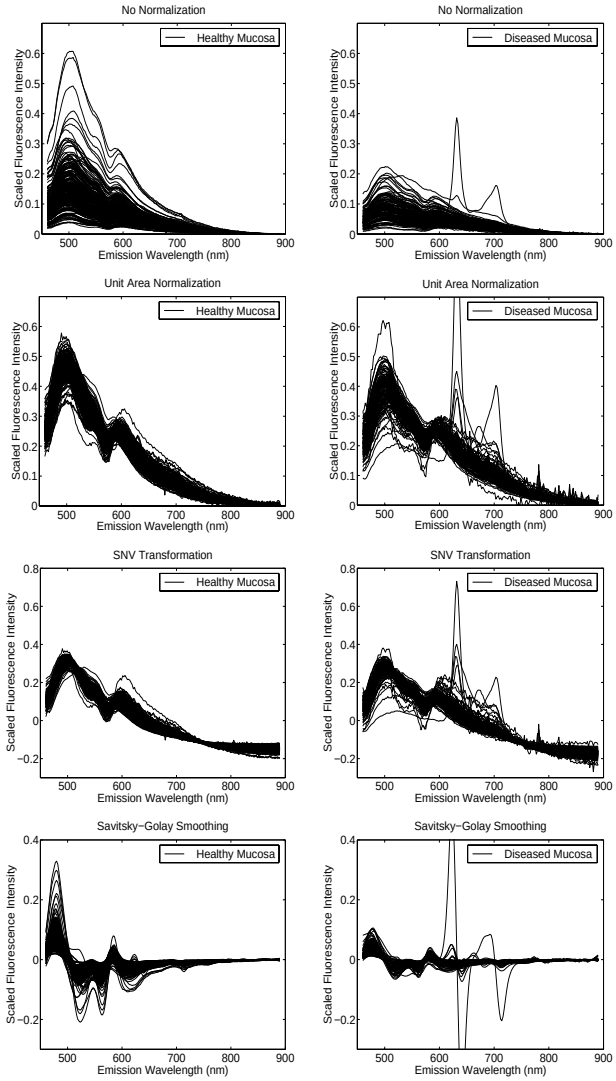

Figure 1. The non-normalized and normalized autofluorescence spectra for healthy and diseased mucosa in oral cavity.

• Savitzky-Golay smoothing and differentiation (SG) [4,5] which is based on performing a least squares linear regression fit of a polynomial of degree k over s data points around each point in the spectrum to smooth the data. We have used the second degree (k=2) polynomial over s=11 bins for smoothing and taken the first derivative of the smoothed spectrum.

Additionally each spectrum (for non-normalized and normalized data) is scaled to be within the interval [-1;1].

Non-normalized and normalized autofluorescence spectra representing healthy and diseased tissues are illustrated in Fig. 1. The medians of autofluorescence spectra for healthy and diseased mucosa are presented in Fig. 2. One can see that different normalization techniques perform differently: dissimilar information may be retrieved from the data. In general, normalization reduces the variance of spectral intensity between spectra belonging to the same data class. However, some information is lost: the distance between the spectral medians may become smaller. This may result in poor lesion diagnostics.

As the measurements at different wavelengths (bins) are strongly correlated, the dimensionality of the problem can be reduced. Such a reduction, in view of the limited number of available samples (we have only 112 spectra representing diseased tissues), can also help to construct a better, more stable classifier [7]. To perform the dimensionality reduction we have applied the principal component analysis (PCA). We have retained 10 leading principal components (that describe 99% of the total variation in the data).

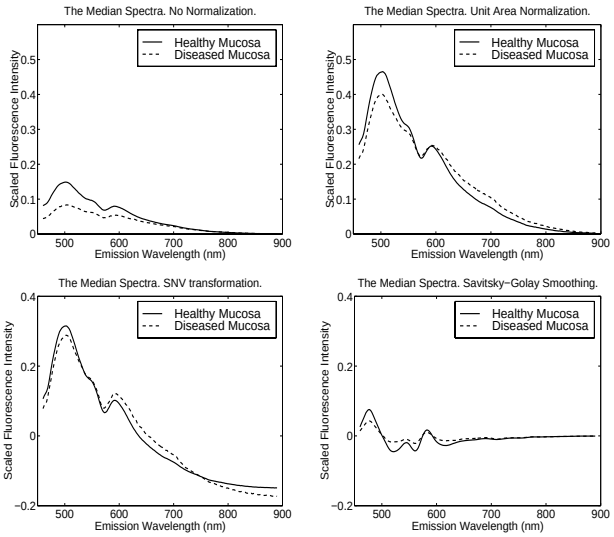

Figure 2. The medians of non-normalized and normalized autofluorescence spectra for healthy and diseased mucosa in oral cavity.

3. Combining Classifiers

Let us now consider the usefulness of combining different normalization techniques in the diagnosis of potentially cancerous lesions. We intend to compare the quality of lesion diagnostics when only one normalization technique (or non-normalized data) is used with a case when all described above normalization techniques are involved in lesion classification. In order to evaluate the quality of lesion diagnostics, we use 2 statistical classifiers. The first one is the regularized Linear Classifier (LC), which constructs a linear discriminant function assuming normal class distributions and using a joint class covariance matrix for both data classes. The second one is the regularized Quadratic Classifier (QC) assuming normal class densities, with regularization parameter $\lambda=10^{-7}$. To obtain our classification results the leave-one-out approach is used. The classification errors (the average of the class errors) of the LC and the QC obtained on leave-one-out for non-normalized and normalized data are presented in the upper part of table 1. The standard deviation of the obtained classification errors is about 0.014.

Table 1: The leave-one-out classification errors (the average of the class errors) for single and combined linear and quadratic classifiers constructed on the first 10 principal components.

	LC	QC
No	0.253	0.253
UA	0.210	0.326
SNV	0.200	0.310
SG	0.246	0.225
wmajority	0.214	0.243
majority	0.248	0.225
product	0.214	0.280
mean	0.214	0.279
max	0.218	0.284
NMC(***L***)	0.203	0.184
NMC(***P***)	**0.156**	**0.174**

We can see that normalizing spectra is useful. But which normalization technique should be preferred over the other ones is a difficult question to answer. For the linear classifier, the unit area normalization and the standard normal variate transformation is a good choice. However, for the quadratic classifier, they both fail. Only Savitzky-Golay smoothing is beneficial. As all normalization techniques transform data in a different way and therefore retrieve different information from the data, it might be worthwhile to use them all for the diagnosis of potentially cancerous lesions.

In order to involve all normalization techniques in lesion classification, we use an approach of combining the classifiers. In this approach, decisions of single classifiers constructed on uniquely normalized data are aggregated into a combined decision. To get a combined decision of all four classifiers constructed on differently normalized data (non-normalized, normalized by unit area, SNV transformation and Savitzky-Golay smoothing), we aggregated the outputs (either obtained labels ***L*** or posterior probabilities ***P***) of these classifiers by several combining rules. We have used five fixed combiners: weighted majority and simple majority voting (applied to labels) [8], and the product, the mean and the maximum combining rules (applied to posterior probabilities) [9]. We have also used one trained combiner - the Nearest Mean Classifier (NMC) [6] (that is almost identical to decision templates method [10]) applied to both labels ***L*** and posterior probabilities ***P*** obtained by the classifiers to be combined. The classification errors (the average of the class errors) of the combined decisions are also estimated using leave-one-out cross-validation technique. They are presented in the lower part of table 1. The standard deviations of the classification errors are between 0.012 and 0.015.

Analysing the results of our simulation study we can see that for our application combining different normalization techniques was not beneficial when fixed combiners were used. Only using a trained combiner on posterior probabilities obtained by single classifiers constructed on uniquely normalized data was useful. We have gained a reasonable improvement for both linear and quadratic classifiers. This occurred because the trained combiner (the nearest mean classifier in our case) retrieves additional information from confidences (the posterior class probabilities) of single classifiers when making a final decision, while fixed combiners are not able to perform this.

4. Conclusions

In lesion classification it is important to get an accurate faultless detection of lesions. This can be challenging if data have a high variance that is caused by measurement techniques and/or variability of measured objects. As a result, the overlap between data classes can be large and lesion diagnostics becomes poor. In order to reduce the variance in data, data normalization can be performed. However, sometimes it is hard to determine the optimum

normalization technique: they all perform equally but the information retrieved by them is different. In this case, it is possible to involve several normalization techniques in lesion classification by using a combined approach when decisions of single classifiers constructed on differently normalized data are aggregated into a final decision. In this case, one can benefit from using the information retrieved by different normalization techniques in lesion diagnostics.

We have tested the approach of combining different normalization techniques for the diagnosis of potentially cancerous lesions performed on autofluorescence spectra measured in the oral cavity. Our simulation study has shown that combining single classifiers obtained on differently normalized data is useful when a trainable aggregating rule is used for making a final decision.

Acknowledgment

This work was supported by the Dutch Cancer Society (Nederlandse Kanker Bestrijding), grant 99-1869, and the Dutch Technology Foundation (STW), grant RRN 5316.

References

1. D.C.G. de Veld, M. Skurichina, M.J.H. Witjes, et.al. Autofluorescence Characteristics of Healthy Oral Mucosa at Different Anatomical Sites, submitted to *Lasers in Surgery and Medicine*, 2003.
2. D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, et.al. *Handbook of Chemometrics and Qualimetrics*. Elsevier, Amsterdam, 1997.
3. A. Gandolfi, R. de Maesschalck, D. Jouan-Rimbaud, P.A. Hailey and D.L. Massart. The Influence of Data Preprocessing in the Pattern Recognition of Excipients Near-Infrared Spectra, *Journal of Pharmaceutical and Biomedical Analysis*, vol. 21, pp. 115-132, 1999.
4. A. Gandolfi, W. Wu, D.L. Massart and S. Heuerding. Comparison of Classification Approaches Applied to NIR-spectra of Clinical Study Lots, *Journal of Pharmaceutical and Biomedical Analysis*, vol. 16, pp. 1329-1347, 1998.
5. P.A. Gorry. General Least-Squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method, *Analytical Chemistry*, vol. 62, pp. 570-573, 1990.
6. K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, San-Diego, 1990.
7. S. Raudys and A.K. Jain. Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 252-264, 1991.
8. Y. Freund, R.E. Schapire. Experiments with a New Boosting Algorithm, *Proceedings of the 13th international conference on Machine Learning*, pp. 325-332, 1996.
9. J. Kittler, M. Hatef, R.P.W. Duin and J. Matas. On T*ransactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226-239, 1998.
10. L.I. Kuncheva, J.C. Bezdek, and R.P.W. Duin, Decision Templates for Multiple Classifier Fusion: An Experimental Comparison, *Pattern Recognition*, vol. 34, no. 2, pp. 299-314, 2001.