Diagnosis of Rheumatoid Arthritis with Optical Tomography: Comparison of Classification Methods

Ludguier D. Montejo[1 *], Julio D. Montejo[2], Hyun K. Kim[1], Uwe J. Netz[4], Christian D. Klose[1], Sabine Blaschke[5], PA Zwaka[6], Gerhard A. Müller[5], Jürgen Beuthan[4], Andreas H. Hielscher[1,3 **]

[1] Dept. of Biomedical Engineering, Columbia University, New York, NY 10025, USA
[2] Dept. of Mathematics, Harvard University, Cambridge, MA 02138, USA
[3] Dept. of Radiology, Columbia University, New York, NY 10025, USA
[4] Laser- und Medizin-Technologie GmbH, Berlin, Germany
[5] Dept. of Nephrology and Rheumatology, Georg-August-University, Göttingen, Germany
[6] Dept. of Radiology, Georg-August-University, Göttingen, Germany
E-mail: (*) ldm2106@columbia.edu, (**) ahh2004@columbia.edu

Abstract—**Linear discriminant analysis (LDA) and support vector machines (SVM) are used to classify reconstructed absorption coefficient distributions of the proximal interphalangeal joints as affected or not affected by rheumatoid arthritis. The performance of each classification method is quantified using the leave-n-out method. LDA is shown to yield high sensitivities, while SVM yields high specificities.**

I. Introduction

Approximately 1.3 million Americans are affected by rheumatoid arthritis (RA) [1]. Patients suffering from RA experience mild to severe inflammation of the synovial of the joints [1,2]. There is no known cure for RA; however, it has been shown that symptoms can be reduced with medication, rest, and exercise if treated during the early stages [2]. Thus, it is crucial to diagnose a patient with RA as early as possible.

Optical tomography (OT) is a promising imaging modality for early diagnosis of RA because it is relatively inexpensive and non-invasive. Differences in optical properties of the synovial in finger joints with and without RA can be visualized in OT reconstructions of the scattering and absorption coefficient distributions [3,4,5]. There are many reasons that account for the change in optical properties of the synovial fluid and surrounding tissue, including an increase in white blood cells. It remains to be determined what is the best possible way to diagnose RA given that the optical properties of joints with RA are significantly altered.

In this paper we build upon efforts previously reported in references [6] and [7] that attempts to correctly diagnose patients with and without RA using classification methods. We classify clinical data using linear discriminant analysis (LDA) and support vector machine (SVM).

II. Methods

A. Data pre-processing

We present results based on clinical data from 158 distinct images; 98 fingers are affected with RA and 60 are not affected. We focus on features from the reconstructed absorption coefficient distribution. The features of interest are the maximum (*max*), minimum (*min*), variance (*var*), and maximum/minimum ratio (*ratio*) across all voxels in the reconstructed distribution.

For each image we extract features only from a region of interest. The region of interest is defined such that the PIP joint is included, only rejecting values within 2mm from the boundary. This is done to avoid boundary artifacts that might be created by the reconstruction process and can potentially dominate the distribution of features of interest [7].

All patients have undergone magnetic resonance imaging (MRI), ultrasound (US), and clinical diagnosis exams, and have subsequently been diagnosed by medical experts and classified as either affected or not affected by RA [4,6,7,8].

B. Classification Methods

The leave-*n*-out method is used to quantitatively assess the performance of each classification algorithm. A subset of data (90%) is used to train, while the remaining data (10%) is used to determine the sensitivity (*Se*), specificity (*Sp*), and Youden index (Y) of the classification method [9].

Each algorithm determines the decision boundary that best separates the data into two classes: affected and not affected. Parameter space is multidimensional; each feature represents one dimension. We present results in 2, 3, and 4 dimensions.

Classification methods vary in the type of decision plane used to separate the two groups. Linear discriminant analysis determines the best line (in 2D) that separates the two groups. The decision on a given data point is made based on its distance from each class centroid [10,11]. SVM differs from LDA in its ability to generate complex decision planes. LDA and SVM are formally treated in [10,12,13] and [10,14] respectively.

III. Results

Fig. 1 is an example of an SVM decision boundary; affected patients are separated from healthy patients. Results at 600MHz using LDA are presented in Table 1. The best

978-1-4244-6923-9/10/$26.00 ©2010 IEEE

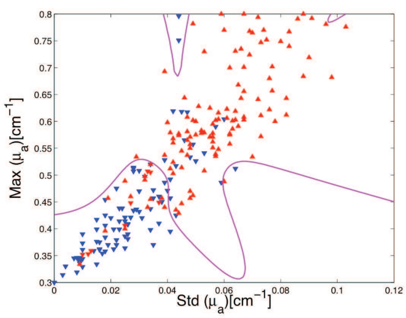

Fig. 1. Typical separation plane using SVM with a sixth degree polynomial kernel. Affected (red, upright triangle) and healthy (blue, upside down triangle) patients are separated by the decision plane.

combination of sensitivity, specificity, and Youden index is {0.85,0.81,0.66} and is given in two-dimensions by *var* and *min*. The largest sensitivity, specificity, and Youden index are 0.93, 0.81, and 0.66, respectively. The best results are obtained in two-dimensions. The lowest classification results also occur in two-dimensions.

Classification results at 600MHz using SVM are presented in Table 2. The best combination of sensitivity, specificity, and Youden index is {0.82, 0.83, 0.64} and is given in two-dimensions by *var* and *ratio*.

Table 1: LDA Classification Results at 600MHz.

Parameter Space	Se	Sp	Y
var, max	0.85	0.75	0.60
var, min	**0.85**	**0.81**	**0.66**
var, ratio	0.80	0.77	0.57
max, ratio	0.83	0.66	0.49
min, ratio	**0.93**	0.62	0.55
var, max, min	0.82	0.76	0.59
var, min, ratio	0.84	0.75	0.59
max, min, ratio	0.89	0.71	0.60
var, max, min, ratio	0.86	0.75	0.62

In general, LDA results have higher sensitivities than results from SVM. The largest sensitivity provided by SVM (0.82) is smaller than all but one sensitivity value provided by LDA. Similarly, SVM results show larger specificity than LDA results. Again, the smallest specificity from SVM is larger than most specificities attained by LDA. The largest single specificity value obtained with SVM is 0.89 and it is given by the features *max* and *ratio*.

We note that the best results are obtained in two-dimensions. However, we also note the lowest classification results also occur in two-dimensions.

Table 2: SVM Classification Results at 600MHz.

Parameter Space	Se	Sp	Y
var, max	0.71	0.80	0.51
var, min	0.80	0.83	0.63
var, ratio	**0.82**	**0.83**	**0.64**
max, ratio	0.72	**0.89**	0.60
min, ratio	0.66	0.87	0.53
var, max, min	0.73	0.80	0.53
var, min, ratio	0.83	0.78	0.61
max, min, ratio	0.76	0.81	0.56
var, max, min, ratio	0.80	0.82	0.62

IV. Conclusion

We have presented classification results of absorption-coefficient distributions in a sagittal cross-section through the PIP joint in 158 individual fingers (98 affected, 60 not affected). The features used for classification are the maximum, minimum, ratio, and standard deviation across all voxels within a region of interest. We have shown that these four parameters show sufficient separation at a modulation frequency of 600 MHz to allow classification using LDA and SVM. The sensitivity is obtained with LDA (0.93), while the largest specificity is obtained with SVM (0.89).

We have shown there is no clear correlation between the number of parameters used for classification (i.e. the dimensionality of the classification method) and success of the classification method. It is also not yet clear if there is a particular feature that works best.

Acknowledgement

This work was supported in part by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS-2R01 AR46255) and the National Cancer Institute (5U54CA126513-03) at the National Institutes of Health

References

1. Arthritis Disease Center, accessed online January 31st, 2010. : http://www.arthritis.org/disease-center.php?disease_id=31.
2. MedicineNet.com, accessed January 31st, 2010. http://www.medicinenet.com/rheumatoid_arthritis/page9.htm.
3. A.H. Hielscher, C.D. Klose, A. Scheel, B. Moa-Anderson, M. Backhaus, U. Netz, J. Beuthan, Sagittal Laser Optical Tomography for Imaging of Rheumatoid Finger Joints, Phys. Med. Bio. 49(7) (2004) 1147-1163.
4. A.K. Scheel, M. Backhaus, A.D. Klose, B. Moa-Anderson, U.J. Netz, K. Hermann, J. Beuthan, G. Müller, G.R. Burmester, A.H. Hielscher, First clinical evaluation of sagittal laser optical tomography for detection of synovitis in arthritic finger joints, Ann Rheum Dis 64 (2005) 239-245.
5. H.K. Kim, A.H. Hielscher, A PDE-constrained SQP algorithm for optical tomography based on the frequency-domain equation of radiative transfer, Inverse Problems 25 (2009) 015010.
6. C.D. Klose, A.D. Klose, U.J. Netz, J. Beuthan, A.H. Hielscher, Optical tomographic detection of rheumatoid arthritis with computer-aided classification schemes, in Multimodal Biomedical Imaging IV, SPIE International Symposium on Biomedical Optics Proc. 7171, paper [7171 11] (2009).
7. U.J. Netz, J. Beuthan, A.H. Hielscher AH, Multipixel system for gigahertz frequency-domain optical imaging of finger joints, Review of Scientific Instruments 79 (2008) 034301.
8. C.D. Klose, A.D. Klose, U. Netz, J, Beuthan, A.H. Hielscher, Multi-parameter classifications of optical tomographic images, Journal of Biomedical Optics 13(05) (2008) 050503.
9. R. Chin, B.Y. Lee, Principles and practice of clinical trial medicine, Elsevier Science & Technology, USA, 2006.
10. S.J. Dixon, R.G. Brereton, Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure, Chemometrics and Intelligent Lab Systems 95 (2009) 1–17.
11. G.J. McLachlan, Discriminant analysis and statistical pattern recognition, New York, Wiley, 1992.
12. W.J. Krzanowski, Principles of Multivariate Analysis, Oxford, Oxford University Press, 1988.
13. N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge, Cambridge University Press, 2000.
14. C. Cortes, V. Vapnik, Support-Vector Networks, Machine Learning 20, 273-297 (1995).