Multiparametric Tissue Abnormality Characterization using Manifold Regularization

Kayhan Batmanghelich[a] and Xiaoying Wu[a] and Evangelia Zacharaki[a] and Clyde E. Markowitz[b] and Christos Davatzikos[a] and Ragini Verma[a]

[a] Section Of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania
[b] Neurology Department, University of Pennsylvania

ABSTRACT

Tissue abnormality characterization is a generalized segmentation problem which aims at determining a continuous score that can be assigned to the tissue which characterizes the extent of tissue deterioration, with completely healthy tissue being one end of the spectrum and fully abnormal tissue such as lesions, being on the other end. Our method is based on the assumptions that there is some tissue that is neither fully healthy or nor completely abnormal but lies in between the two in terms of abnormality; and that the voxel-wise score of tissue abnormality lies on a spatially and temporally smooth manifold of abnormality. Unlike in a pure classification problem which associates an independent label with each voxel without considering correlation with neighbors, or an absolute clustering problem which does not consider a priori knowledge of tissue type, we assume that diseased and healthy tissue lie on a manifold that encompasses the healthy tissue and diseased tissue, stretching from one to the other. We propose a semi-supervised method for determining such as abnormality manifold, using multi-parametric features incorporated into a support vector machine framework in combination with manifold regularization. We apply the framework towards the characterization of tissue abnormality to brains of multiple sclerosis patients.

Keywords: Tissue Abnormality Characterization, Manifold Regularization, Support Vector Machine, Multiple sclerosis, Lesions, Normal Appearing Brain Tissue

1. INTRODUCTION

In order to obtain a more comprehensive characterization of tissue abnormality, combination of information from several MR protocols, has gained attention recently,[1,2] While this has mainly been employed to detect diseased tissue such as lesions, but it is expected that it may also help in quantifying the deviation of tissue from *healthiness* that is, the degree of abnormality.[3] Tissue characterization of abnormality helps identify regions that are being progressively affected by disease and as such can be used for treatment planning and prognosis. In addition, tissue characterization, if well validated, could also be used in the form of a segmentation tool. However, abnormality manifests itself differently in different magnetic resonance imaging (MRI) protocols. Hence there is a need for sophisticated statistical methods that combine information from the different protocols to provide a comprehensive tissue abnormality profile.

There are very few publications that deal with tissue characterization with the majority concentrating on segmentation of diseased tissue (lesion/tumor). Welti et. al[4] tried to characterize MS lesions by segmenting them into active lesions by analyzing the spatio-temporal behavior of tissue voxels. But they do not characterize the degree of abnormality in the white matter outside of the lesions that is also deteriorating as part of the progression of disease. A course adopted via discriminative methods uses samples of various classes to train models for each tissue class and provides good discriminative power in terms of class separation. Support Vector Machines (SVM),[5] Bayesian,[6] or Neural networks,[7] are usually used as the core of such methods. These methods treat each voxel individually with no spatial constraints with neighboring voxels incorporated. These discriminative models require a lot of samples from each class in order to give accurate result. Obtaining reliable

Further author information: (Send correspondence to K.B.)
K.B.: E-mail: nematollah.batmanghelich@uphs.upenn.edu, Telephone: 1 215 349 8573
R.V.: E-mail: Ragini.Verma@uphs.upenn.edu, Telephone: 1 215 662 7471

Medical Imaging 2008: Computer-Aided Diagnosis, edited by Maryellen L. Giger, Nico Karssemeijer
Proc. of SPIE Vol. 6915, 691516, (2008) · 1605-7422/08/$18 · doi: 10.1117/12.770837

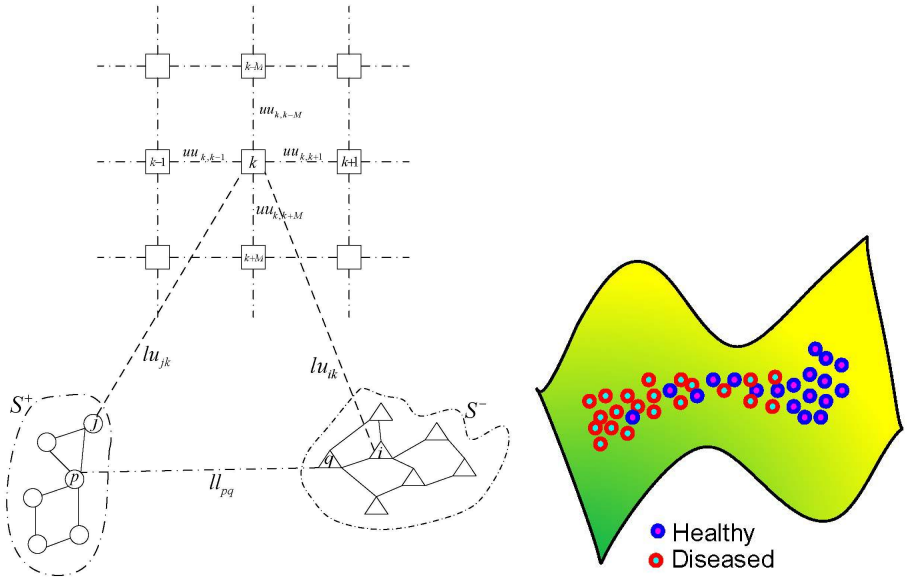

Figure 1. **Left**: Semi-supervised learning: bag of positive S^+ denoting healthy and negative examples S^- depicting diseased samples are available. Grid represents voxels that are not labeled but lie in between healthy and diseased in terms of abnormality. **Right**: Distribution of samples that vary on a smooth manifold encompassing the normal cluster at one end to the abnormal cluster at the other.

ground truth usually requires an expert to go through tedious manual segmentation. Corso et al.[8] proposed SWA (Segmentation by Weighted Aggravation) which is implemented within a Bayesian framework with hierarchical weight assignment. Song et al.[9] suggested semi-automated tissue segmentation using Gaussian Process. In all these cases, segmentation or in other words, separation of two classes was the main interest, but not a continuous characterization of abnormality, as we provide in our framework.

In this paper, we consider tissue abnormality characterization as a regression problem. We assume that there is a smooth manifold encompassing normal and lesion tissue with normal appearing abnormal tissue lying on this manifold. Furthermore, abnormality score should be continuous and smooth on manifold and spatially on the image. However, there are some issues to be addressed: first, conventional regression methods like Support Vector Regression (SVR) does not provide leverage to control the smoothness; second, abnormality in the absence of consistent ground truth may be easier to be characterized using combination of MR protocols rather than a single one. We have some samples from healthy brain and some samples from lesion part of diseased brain as labeled training samples. Voxels of the brain which are to be tested are considered as unlabeled samples. Taking advantage of Laplacian Regularized Least Square (LapRLS)[10] formulation as a semi-supervised regression method, we associate a continuous abnormality score pertaining to each voxel of the brain implemented as an embedding graph consisting of labeled and unlabeled voxels. Training samples and unlabeled voxels set up vertices of an embedding graph. Associations between neighborhood voxels are taken care of using a proper edge weighting scheme between vertices of embedding graph. Since a smoothness constraint is imposed on the cost function of regression, the result of such a functional optimization could be treated as a qualifier of tissue which provides the abnormality characterization. Employing LapRLS, we propose a method which can handle both the criteria (spatial and manifold smoothness) in one framework, so that a continuous abnormality score is obtained. The framework is applied to multi-parametric data acquired on MS patients with the idea of characterizing not only the lesions diseased or healthy tissue, but also the WM that is progressing to abnormality based on the stage of the disease.

2. FRAMEWORK FOR TISSUE ABNORMALITY CHARACTERIZATION

Our reasons for using manifold regularization method for tissue characterization is two folds: first of all, we want to address the fact that samples that are close to each other should possess a similar decision function. The

meaning of closeness is defined based on graph embedding and edge weights which have been defined between vertices. For example, if two voxels of an unlabeled image are close together spatially, they should be associated with a similar decision function. As an alternative example, if an individual voxel is similar to one of the labeled training samples, it should in high probability have the same decision value as the labeled instance (Fig.1:Left). Secondly, as we are interested in tissue abnormality characterization, assigning discrete numbers (class labels) is not sufficient because it is not depictive of the gradual transition of tissue from diseased to healthy. In other words, they are located somewhere between healthy and diseased tissue samples on a manifold (Fig.1:Right). In addition, the proposed framework is capable of addressing and incorporating the relationship between adjacent voxels in image (Fig.1). In such a case, a smooth decision function is required. To do so, we have assumed that healthy and diseased tissue lie on two disjoint clusters however tissue with intermediate degrees of abnormality fill the gap in between two clusters. However, we do not have any ground truth for this type of intermediate tissue. Therefore, we adopt a manifold regularization[10] framework to address it. To do so, we have assumed that healthy and diseased tissue lie on two disjoint distributions however tissue with intermediate degrees of abnormality fill the gap in between two distributions.

2.1 Laplacian Regularized Least Square

Given a set of labeled example $(x_i, y_i), i = 1, ..., l$, (in our case, x_i's are voxels with multi-parametric intensities that have been labeled (y_i) as diseased or healthy), our aim is to find a function, namely abnormality function f, which satisfies following condition:

$$f^* = \arg\min_{f \in \mathcal{H}_K} \frac{1}{l} \sum_{i=1}^{l} \mathcal{C}(x_i, y_i, f) + \lambda_{\mathcal{R}} \|f\|_K^2 \ + \lambda_{\mathcal{M}} \|f\|_{\mathcal{M}}^2 \tag{1}$$

The first term, $\mathcal{C}(x_i, y_i, f)$, penalizes error for labeled samples, y_i, which in the case of regression, it could be square loss function $\mathcal{C}(x_i, y_i, f) = (y_i - f(x_i))^2$. The second term, $\lambda_{\mathcal{R}} \|f\|_K^2$, and third term, $\lambda_{\mathcal{M}} \|f\|_{\mathcal{M}}^2$, together impose different smoothness conditions on the abnormality function. First one imposes smoothness such that normal and lesion samples would not be mixed together. The latter one takes care of spatial smoothness and smoothness of abnormality score between labeled and unlabeled samples. f^* is abnormality score derived from minimization of eq.1 in which $f^* > 0$ is considered abnormal and $f^* < 0$ is considered normal. With some reasonable mathematical assumptions described in,[10] the last term $\|f\|_{\mathcal{M}}^2$,can be approximated by graph Laplacian which is constructed based on labeled and unlabeled samples. As it is shown in,[10] the optimized function will be:

$$f^* = \arg\min_{f \in \mathcal{H}_K} \frac{1}{l} \sum_{i=1}^{l} \mathcal{C}(x_i, y_i, f) + \lambda_{\mathcal{R}} \|f\|_K^2 \ + \frac{\lambda_{\mathcal{M}}}{(u+l)^2} \hat{f}^T L \hat{f} \tag{2}$$

where $\hat{f}$ is a vector containing outcome (class label) for labeled and unlabeled samples. Since we do not have outcome for unlabeled samples, corresponding elements in $\hat{f}$ will be zeros. Matrix $L = D - W$ is a graph Laplacian matrix and W matrix contains edge weights of embedding graph and D is a diagonal matrix diagonal elements of which are sum of corresponding columns.[10] As it is shown in[10] the decision function would be in the form of $f(x) = \sum_{i=1}^{u+l} \alpha_i K(x_i, x)$. In this study, we have used the RBF kernel.

As the square loss function, $(y_i - f(x_i))^2$, has been used, the resulting optimization problem would be Laplacian Regularized Least squared (LapRLS) which is a type of regression problem. In that case, optimal solution for α_i could be derived from following linear system:[10]

$$\alpha^* = (JK + \lambda_{\mathcal{R}} lI + \frac{\lambda_{\mathcal{M}} l}{(u+l)^2} LK)^{-1} Y \tag{3}$$

Where Y is a vector containing labels for trainig samples (+1,-1,0 for lesion, healthy and unlabeled samples respectively). $J = diag(y_1, y_2, ..., y_l, 0, 0, ..., 0)$ is a diagonal matrix holding labels for labeled samples and zero for unlabeled samples on diagonal elements. I is an identity matrix.

Unlike traditional transductive SVM method like[11] in which similarity between labeled and unlabeled samples cannot be defined arbitrarily, the interesting point of manifold-based methods is that they give us a tool to

define similarity between samples arbitrarily. In fact, prior knowledge could be incorporated in terms of weights between samples, such as in You et al.[12] and Geng et al.[13] Since the main goal of this paper is not classification (lesion segmentation), but aims at tissue characterization, we should design a weighting scheme which can reveal underlying abnormality coded by various MR modalities. As it is mentioned earlier, we have asssumed that abnormal tissue samples are scattered between normal and lesion clusters on the sample manifold. In addition, voxels which are spatially close should supposedly possess similar abnormality values. Taking into account these two facts, we propose the weighting scheme which helps us characterize tissue while imposing spatial and features space smoothness, simultaneously.

In this scheme, namely *Additive weighting scheme*, edge weights are assigned with the following rules:

$$L = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22}^{+} \end{bmatrix} \tag{4}$$

$$L_{12} = L_{21}^{T}$$

$$L_{22}^{+} : [l_{22}^{+}]_{ij} = \eta [l_{22}]_{ij} + (1-\eta) e^{\left(\frac{-(x_i^p - x_j^p)^2}{2\sigma^2}\right)}$$

in which L_{11}, L_{12}, L_{21} are ordinary weight, e.g. binary nearest neighbor weights or heat kernel or any other method, which is solely derived from features, in our case multi-parametric intensity features. L_{11} in a block matrix represents the weights between labeled samples and L_{21} and L_{21} respectively shows the weights between labeled-unlabeled and unlabeled-labeled samples respectively, and L_{22} corresponds to unlabeled-unlabeled samples weights. x_i^p and x_j^p are positions vectors of i'th and j'th voxels in space respectively. Equation (4) says that if there is any relationship, say nearest neighbor, between samples of labeled instances, it should be kept intact; for unlabeled voxels, Eq.4 combines weights derived from features L_{22} and their location with each other with arbitrary ratio η to derive new weights L_{22}^{+}. In fact, η and σ could be seen as leverages to control spatial smoothness. Mathematical details about how to derive Eq.3 and more can be found in[10] .

Tissue Abnormality Profile: On applying LapRLS to a dataset, the abnormality score produces a voxel-wise smooth measure of abnormality which is referred to as the Tissue Abnormality Profile. This is color coded to visualize the level of abnormality. Red indicates high abnormality and blue indicates healthy. These tissue abnormality maps can be segmented (using a threshold) to obtain maps of diseased tissue, that is, lesions. These regions are depicted in red in Fig.2:**(c),(d)**. In all color maps, minimum and maximum of abnormality score values are normalized to zero and one respectively.

3. EXPERIMENTAL DESIGN AND RESULTS

We have applied our framework to a dataset of patients with multiple sclerosis (MS). The aim is to be able to characterize the abnormality in Normal Appearing Brain Tissue (NABT) in the patients, using samples identified in the lesions in the patients and the healthy tissue in the controls. Multiparametric intensity feature was designed at each voxel incorporating the intensities from different MR protocols. In each case several MR protocols were acquired: namely, FLAIR, T1, T2, Magnetization Transfer (MT), and diffusion tensor imaging(DTI). Fraction Anisotropy(FA) and Trace maps are calculated from the DTI data, and B0 (image acquired without gradient direction) are being used. For training samples for healthy tissue, various parts of the brain of the healthy controls were outlined which should include gray and white matter in addition to CSF. Multiparametric Intensity features are defined at each voxel, by concatenating the intensity value at that voxel from each protocol into a vector. To show our method can capture abnormality and due to the fact that there is no ground truth for NABT, we will study the behavior of abnormality over time (Fig.2:**(c),(d)**).

Fig.2:**(d)** shows how normal appearing brain tissue that is progressing to abnormal evolves over time and converts to purely lesion tissue. The time interval between successive scans are about three months. However, MS lesions may appear and disappear in less than two months; in other words, it is possible that a part of tissue that appears purely healthy in the first scan progressively converts to abnormal and eventually to lesion. As it can be seen in Fig.2:**(d)**, although a very diffuse white area is evident in the peri-ventricular region in the first time point(T0), with no high intensity lesion, however our decision function visualization is able to pick it

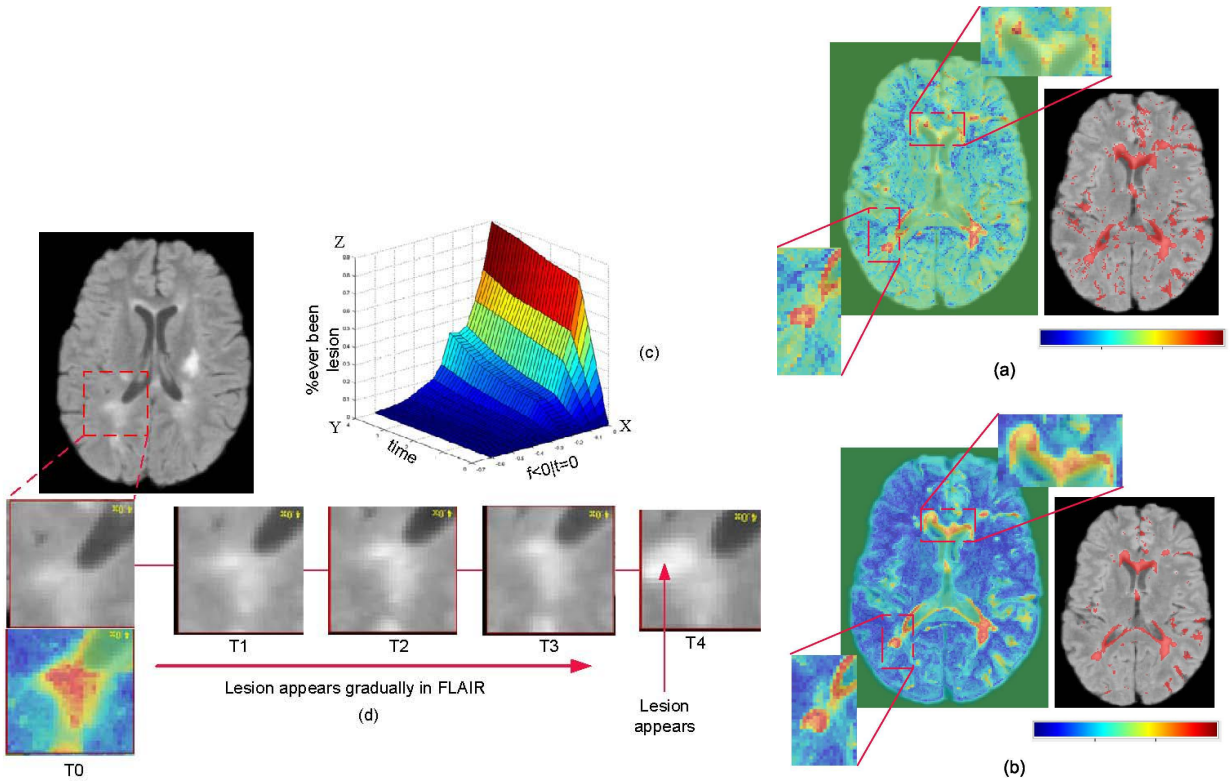

Figure 2. **(d)**: LapRLS result for peri-ventricular area on first time point: it becomes high intense in FLAIR modality which is the manifestation of MS lesion. Eventually, it becomes a full-fledged lesion. **a**: shows how the abnormality score identifies abnormality in areas which are *classified* as healthy. It shows that the portion of tissue with high abnormality value (f) becomes lesion over time rather than those which posses lower abnormality value. Z-axes represents the chance of tissue to ever converting to lesion until $t = T_i$. **(a),(b)**: These are the tissue abnormality map. Effect of manifold and spatial smoothing. Each of the figures on the left show the color-coded voxel-wise outcome of the decision function on applying manifold and spatial smoothing. The figure on the right, is obtained by threholding the tissue abnormality map to determine the regions of high abnormality (or lesions). (a)Less weight for manifold smoothing (last term two terms in Eq.1) terms lead to less smooth decision function and noisier labeling as indicated by the patchy color map; This shows that both manifold($\lambda_{\mathcal{M}}$) and spatial smoothing (η) are important for obtaining a smooth tissue characterization (and hence a smoother tissue abnormality map). The more reddish, the more abnormal tissue would be and the more bluish, the healthier tissue is.

up as abnormal in the color map. The region progressively becomes higher intensity in the FLAIR modality which is characteristic of MS lesion and eventually, it becomes a full-fledged lesion(T4). Fig.2:**(c)** shows that the more abnormal the tissue is initially, the more likely it is to convert to lesion over time. X-axis in Fig.2:**(c)** shows the outcome of the abnormality score(f) on the first time point only for tissues which are *classified* as healthy($f < 0$); Y-axis shows *time* point and Z-axis shows the likelihood of ever been lesion. It shows that tissues, abnormality score of which are closer to the boundary separating the healthy from the boundary ($f = 0$), are more likely to eventually become lesion rather those which are further away from boundary. Fig.2:**(a),(b)** shows that decreasing the effect of the manifold terms, lower values for $\lambda_{\mathcal{R}}$ and $\lambda_{\mathcal{M}}$, (second and third terms in Eq.1), results in a noisier voxel-wise tissue label and abnormality map. The combined effect of the parameters yields a smoother result indicating the need for both spatial and manifold smoothness.

4. DISCUSSION

Our framework incorporates manifold regularization to classify voxels with multi-modal features. The outcome of application to real data is to obtain a gradually varying decision value (that is, a smoother color map visualization)

based on expert defined labeled voxels. Although many have applied SVM or other classification techniques to classify each voxel individually and independently, our method takes into account the spatial position of samples and their arrangement on data manifold simultaneously. Therefore, as is expected, it yields results which are smoother and consequently smoothness removes some false positives (identifying voxels as lesions or abnormal when they are not). In addition, we obtain classification even in the presence of very subtle differences in tissue characterization. Thus our method is able to capture the manifold of disease progression determined by the healthy tissue at one end and lesion tissue at the other. Through the process of employing manifold regularization, we are able to capture smooth transistion from the normal, to normal-appearing and finally to fully diseased tissue. Owing to the smooth transistion and effective classification, the framework is potentially better for lesion segmentation also, in addition to the characterization of tissue abnormality. The framework is particularly applicable to high dimensional data which is difficult to regularize in general, and obtain a framework for tissue abnormality characterization.

5. CONCLUSION AND FUTURE WORK

In this paper we have proposed a framework for characterizing the tissue abnormality by treating abnormality to lie the healthy and diseased tissue on a manifold. The application of the manifold regularization methods to the multi-parametric MS datasets, produce a voxel-wise decision map of abnormality that can be used to investigate tissue that is progressively becoming abnormal. We expect that such a novel characterization of tissue abnormality will help determine tissue that is deteriorating and help in prognosis and determining treatment, thereby changing the course of the disease.

REFERENCES

1. Y. Wu, S. K. Warfield, and et al., "Automated segmentation of multiple sclerosis lesion subtypes with multichannel mri," *NeuroImage* **32**(3), pp. 1205–1215, 2006.
2. P. Anbeek, K. L. Vincken, and et al., "Automatic segmentation of different-sized white matter lesions by voxel probability estimation," *Med. Imag. Anal.* **8**(3), pp. 205–215, 2004.
3. D. S. Meier and C. Guttmann, "Time-series analysis of mri intensity patterns in multiple sclerosis," *NeuroImage* **20**(2), pp. 1193–1209, 2003.
4. D. Welti, G. Gerig, and et al., "Spatio-temporal segmentation of active multiple sclerosis lesions in serial mri data," in *IPMI'01*, pp. 438–445, 2001.
5. R. J. Ferrari, X. Wei, and et al., "Segmentation of multiple sclerosis lesions using support vector machines," in *SPIE'03*, **5032**, pp. 16–26, 2003.
6. G. D.-Phocion, M. Gonzalez, and et al., "Hierarchical segmentation of multiple sclerosis lesions in multi-sequence mri," in *ISBI'04*, pp. 157–160, 2004.
7. A. Hadjiprocopis and P. Tofts, "An automatic lesion segmentation method for fast spin echo magnetic resonance images using an ensemble of neural networks," in *IEEE Workshop on Neur. Net. for Sig. Proc., NNSP'03*, pp. 709–718, 2003.
8. J. J. Corso, E. Sharon, and et al., "Multilevel segmentation and integrated bayesggytian model classification with an application to brain tumor segmentation," in *MICCAI'06*, **2**, pp. 790–798, 2006.
9. Y. Song and et al., "A discriminative method for semi-automated tumorous tissue segmentation of mr brain images," in *CVPR'06*, pp. 79–86, 2006.
10. M. Belkin, P. Niyogi, and V. Sindhwani, "On manifold regularization," in *Workshop on AISTAT'05*, 2005.
11. T. Joachims, "Transductive inference for text classification using support vector machines," in *ICML'99*, 1999.
12. Q. You and et al., "Neighborhood discriminant projection for face recognition," in *ICPR'06*, **2**, pp. 532–535, 2006.
13. X. Geng and et al., "Supervised nonlinear dimensionality reduction for visualization and classification," *IEEE Trans. on Sys., Man, and Cyber-Part B:Cyber.* **35**(6), pp. 1098–1107, 2005.