SEGMENTING HUMAN KNEE CARTILAGE AUTOMATICALLY FROM MULTI-CONTRAST MR IMAGES USING SUPPORT VECTOR MACHINES AND DISCRIMINATIVE RANDOM FIELDS

Kunlei Zhang, Jun Deng, Wenmiao Lu
School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore

ABSTRACT

This paper presents a novel solution toward the accurate and automatic cartilage segmentation with multi-contrast MR images based on pixel classification. The previous pixel classification based works for cartilage segmentation only rely on the labeling by a trained classifier, such as support vector machines (SVM) or k-nearest neighbors. However, these frameworks do not consider the spatial information. To incorporate spatial dependencies in pixel classification, we explore a principled framework of pixel classification based on the convex optimization of an SVM-based association potential and a discriminative random fields (DRF) based interaction potential for our task of cartilage segmentation. The local image structure based features as well as the features based on geometrical information are adopted as the features. We finally perform the loopy belief propagation inference algorithm to find the optimal label configuration. Our framework is validated on a dataset of multi-contrast MR images. Experimental results show that the combined features compare favorably to the two types of separate features and our pixel classification framework outperforms the conventional frameworks based solely on SVM or DRF for cartilage segmentation in subject-specific training scenario.

Index Terms—Automatic segmentation, discriminative random fields, support vector machines, MRI, cartilage

1. INTRODUCTION

Human knee joint is commonly affected by osteoarthritis (OA), a prevalent disease mainly characterized by the degradation of the articular cartilage. Since the assessment of OA requires exact quantification of the cartilage, accurate cartilage segmentation as the key to such quantification has gained considerable attention in recent years. Manual and semi-automatic segmentation are typically labor intensive with reconstruction time up to several hours and prone to inter- and intra-observer variability. It is thus advantageous to automate the segmentation. However, automatic cartilage segmentation from MR images is a very challenging task due to many factors, such as the thin variable morphology of the cartilage, variable intensity homogeneity, the low contrast between the cartilage and other soft tissues, and MR artifacts. In this paper, we address the problem of automatic cartilage segmentation with multiple sets of MR images taken with different sequences (referred to as *multi-contrast MR images*). To the best of our knowledge, the research on cartilage segmentation from multi-contrast MR data is sparse.

1.1. Related Works

Several works automatically segmented cartilages from MR images using atlas [1] or deformable-model [2, 3] based approaches which typically require a large training datasets to get the prior knowledge. Without formulating the prior knowledge, image segmentation can also be considered as a statistical classification problem in which each pixel belongs to a class. Folkesson et al. [4] proposed to use a two step k-nearest neighbors (k-NN) classifier to automatically separate cartilages from non-cartilages. Each pixel was described by selected local image structure based features. While all the aforementioned works have performed cartilage segmentation from a single MR sequence, multi-contrast MR images can provide different contrast mechanisms between tissues and would help separate different tissues. Koo et al. [5] proposed to segment cartilage automatically with multi-contrast MR data using the SVM, and utilized intensities as well as geometrical information as the features. However, these pixel classification based works [4, 5] assumed that individual pixels were independent, which may not be appropriate for the cartilage segmentation task. As pointed out by [6], class labels are not independent in most real-world spatial classification problems, where correlations in the labels exist in data with multi-dimensional structure, such as images and volumes. This motivates us to incorporate spatial dependencies in the pixel classification for cartilage segmentation.

1.2. Overview of the Work Presented

In this paper, we present a novel solution to the problem of automatic cartilage segmentation with multi-contrast MR images based on SVM and DRF, which not only incorporates the spatial dependencies in pixel classification but benefits from the generalization ability of the classification based on SVM. Specifically, we adopt the pixel classification framework based on the minimization of the convex combination of the SVM-based association potential and the DRF-based interaction potential for cartilage segmentation. We also employ a feature set encoding diverse forms of image and anatomical structure information, including intensities, image local structure based features [4] and geometrical information [5] all of which are crucial and helpful to distinguish cartilages from other tissues. The block diagram of the presented segmentation framework is shown in Figure 1. Although the framework of the combination of SVM and DRF has been proposed for prostate cancer localization [7] and brain tumor extraction [8], it has not been specialized for cartilage segmentation. Unlike these works, we perform the loopy belief propagation (LBP) for the inference in this paper. The major contributions of this paper lie in the following. First, this is the first work of incorporating the spatial dependencies in the pixel classification for cartilage segmentation. Second, the image local structure based features and the features based on geometrical information are combined as the features in our framework, which outperforms the pixel classification based on either of the two types of features.

2. METHODOLOGY

2.1 Description of Multi-contrast Human Knee MR Images and Preprocessing

We use a database consisting of multi-contrast human knee MR

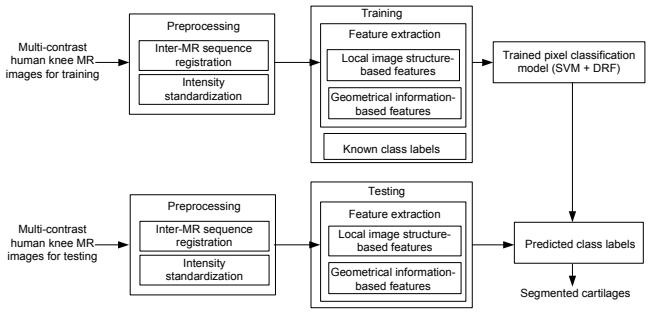

Figure 1. The block diagram of the presented segmentation framework

images from [4]. Multiple sets of MR images were taken for a knee joint from each subject using a 3.0T magnet scanner (GE Healthcare, Waukesha, WI) with multiple MR sequences including FS SPGR, FIESTA, and IDEAL GRE (water and fat images). Different MR sequences were acquired using different sets of parameters. FS SPGR sequences were obtained with echo time (TE) 4ms, repetition time (TR) 28ms, and flip angle (FA) 25°; FIESTA with TE 3.5ms, TR 7.2ms and FA 25°; IDEAL GRE (water & fat) with TE 3.4ms, TR 9.4ms and FA 30°. In addition, several parameters were common, with all images acquired in the sagittal plane with slice thickness 1.5mm, in-plane spacing 0.625×0.625 mm^2 and matrix size 256×256. The golden standard cartilage segmentation for each knee was obtained by an expert's manual segmentation in the form of binary mask images of cartilage.

Preprocessing is required before segmenting cartilage from these knee MR data. The multiple sets of MR images are aligned using the MultiResMI Registration in ITK [9]. We removed regions that contain no cartilage information and the volume size is reduced to $150\times188\times55$. Intensities of all sets of MR images were normalized to [0 1], which brings them within the same dynamic range, improving the stability of the classifier.

2.2. Segmentation Methods

Throughout this paper, let $\mathbf{y}=\{\mathbf{y}_i\}_{i\in S}$ represent the observed data, where the observation $\mathbf{y}_i$ of the ith site is typically represented by d-dimensional feature vector, and S is the set of sites. Let the corresponding labels at the image sites be given by $\mathbf{x}=\{x_i\}_{i\in S}$, where x_i is the class label of the ith site. The aim of the classification based segmentation is to infer the most likely joint class labels $\mathbf{x}^*=\{x_i^*\}_{i\in S}$ based on the observed data $\mathbf{y}$ by a classification model.

2.2.1 Classification Model

Although SVM has been used to segment cartilages in [5], it does not consider spatial information in the classification process, which is necessary and beneficial to the cartilage segmentation. Probabilistic graphical models such as Markov random fields (MRF) and DRF have been used to incorporate spatial contextual constraints in many applications [6]. In our cartilage segmentation task, one of the main challenges is that anatomic structures in particular the cartilages have complex shapes and variable appearances in MR images, which causes that the observed MR data may not be appropriately modeled by some specifically factorized forms. Since DRF alleviates the need to model the distribution over the observations, we thus select DRF instead of the generative MRF to incorporate spatial dependencies for cartilage segmentation. In DRF model [6], the joint distribution over the labels $\mathbf{x}$ given the observations $\mathbf{y}$ can be written as,

$$p(\mathbf{x}\mid\mathbf{y})=\frac{1}{Z}\prod_{i\in S}A(x_i,\mathbf{y})\prod_{i\in S}\prod_{j\in N_i}I(x_i,x_j,\mathbf{y}) \tag{1}$$

where Z is a normalization constant called the partition function, $A(x_i,\mathbf{y})$ is the association potential modeling dependencies between the label x_i and the observations $\mathbf{y}$, and $I(x_i,x_j,\mathbf{y})$ is the pairwise interaction potential modeling dependencies between the labels x_i,x_j and the observation $\mathbf{y}$. A generalized linear model (GLM) with the logistic function as a link is used for $A(x_i,\mathbf{y})$ [6],

$$A(x_i,\mathbf{y})=1/(1+\mathrm{e}^{-x_i\mathbf{w}^T\mathbf{f}_i(\mathbf{y})}) \tag{2}$$

where $\mathbf{w}$ is the model parameter, $\mathbf{f}_i(\mathbf{y})$ is the feature vector for node i. The interaction potential is modeled as,

$$I(x_i,x_j,\mathbf{y})=\mathrm{e}^{x_ix_j\mathbf{v}^T\mathbf{f}_{ij}(\mathbf{y})} \tag{3}$$

where $\mathbf{v}$ is the model parameter, $\mathbf{f}_{ij}(\mathbf{y})$ is the feature vector for nodes i and j computed from the observations $\mathbf{y}$. $\mathbf{f}_{ij}(\mathbf{y})$ can be set by taking the absolute difference of $\mathbf{f}_i(\mathbf{y})$ and $\mathbf{f}_j(\mathbf{y})$, that is, $\mathbf{f}_{ij}(\mathbf{y})=|\mathbf{f}_i(\mathbf{y})-\mathbf{f}_j(\mathbf{y})|$ which penalizes for high absolute differences in the features of neighboring nodes.

To improve upon the cartilage segmentation based on SVM in [5], we adopt the unified formulation of SVM and DRF in [8] which was originally applied to brain tumor extraction. In addition, the unified formulation of SVM and DRF is also preferable to the original DRF based on GLM because of the appealing generalization ability of SVM. Therefore, as in [8], we utilize SVM in the association potential. Then, the unified model can be written as

$$p(\mathbf{x}\mid\mathbf{y})=\frac{1}{Z}\prod_{i\in S}\frac{1}{1+\mathrm{e}^{-x_i(u_0+u_1\gamma(\mathbf{f}_i(\mathbf{y})))}}\prod_{i\in S}\prod_{j\in N_i}\mathrm{e}^{x_ix_j\mathbf{v}^T\mathbf{f}_{ij}(\mathbf{y})} \tag{4}$$

where $\gamma(\bullet)$ in the association potential is the SVM decision function which can be described as [10]

$$\gamma(\mathbf{y})=\sum_{l\in L}\alpha_l x_l k(\mathbf{y}_l,\mathbf{y})+b \tag{5}$$

where the parameter α_l and a bias b can be learned through solving a quadratic programming problem [10], $\mathbf{y}_l$ denotes support vector, L is the set of indices of the support vectors, and $k(\bullet)$ is the kernel function that can map the observation into higher dimensionality feature space. The most commonly used kernel for classification is the Gaussian radial basis function with $k(\mathbf{y}_i,\mathbf{y}_j)=\exp(^{-\|\mathbf{y}_i-\mathbf{y}_j\|}/_{2\sigma^2})$. In the model (4), $\mathbf{u}=[u_0,u_1]^T$ and $\mathbf{v}$ are the association potential parameters and interaction potential parameters, respectively, to be estimated in the training stage. The features used in this work are to be determined in the following subsection.

2.2.2. Features

Features based solely on intensity are insufficient to extract cartilage due to overlapping intensity distributions and ambiguous boundaries between the cartilage and non-cartilages. Hence, we also adopt the local image structure related features including Gaussian smoothed intensities, the first- and second-order derivatives, and three eigenvalues of the Hessian, which have been successfully used for cartilage segmentation task in [4]. Further, the geometrical information of the anatomical structure in the multi-contrast MR images has proven to be crucial to distinguish cartilage from surrounding tissues such as muscle [5]. The geometrical information for each pixel used in this paper include

Euclidean distance from the closest bone and gradient of the distance, angle between the main magnetic direction of the MRI and the line to the femoral center line, and relative location along the medial and lateral center of the distal femur.

Thus, the feature set adopted in this paper consists of normalized intensities of multi-contrast MR images, features related local structure of images, and geometrical information based features.

2.2.3. Learning and Inference

In the classification model (4), we first compute the SVM decision function $\gamma(\bullet)$ by solving a quadratic programming problem using SVMlight [11]. Then the parameters $\{\mathbf{u}, \mathbf{v}\}$ are simultaneously estimated with M training images using pseudolikelihood,

$$(\hat{\mathbf{u}}, \hat{\mathbf{v}}) = \arg\max_{(\mathbf{u},\mathbf{v})} \prod_{m=1}^{M} \prod_{i \in S} p(x_i^m \mid x_{N_i}^m, \mathbf{y}^m, \mathbf{u}, \mathbf{v}) \qquad (6)$$

where

$$p(x_i^m \mid x_{N_i}^m, \mathbf{y}^m, \mathbf{u}, \mathbf{v}) = \frac{1}{z_i^m} e^{x_i^m \mathbf{u}^T \boldsymbol{\gamma}(\mathbf{f}_i(\mathbf{y}^m))} \prod_{j \in N_i} e^{x_i^m x_j^m \mathbf{v}^T \mathbf{f}_{ij}(\mathbf{y}^m)}$$

with

$$\boldsymbol{\gamma}(\mathbf{f}_i(\mathbf{y}^m)) = [1 \quad \gamma(\mathbf{f}_i(\mathbf{y}^m))]^T$$

and

$$z_i^m = \sum_{x_i^m \in \{-1,1\}} \left(e^{x_i^m \mathbf{u}^T \boldsymbol{\gamma}(\mathbf{f}_i(\mathbf{y}^m))} \prod_{j \in N_i} e^{x_i^m x_j^m \mathbf{v}^T \mathbf{f}_{ij}(\mathbf{y}^m)} \right)$$

To prevent over-fitting, we utilize the L2-regularization, which involves adding a penalty term with the form of a sum of squares of all the parameters in order to discourage the parameters from reaching large values. We thus have the following penalized negative logarithm pseudolikelihood

$$(\hat{\mathbf{u}}, \hat{\mathbf{v}}) = \arg\min_{(\mathbf{u},\mathbf{v})} \Big[\sum_{m=1}^{M} \sum_{i \in S} \big(\log(z_i^m) - x_i^m \mathbf{u}^T \boldsymbol{\gamma}(\mathbf{f}_i(\mathbf{y}^m)) - \sum_{j \in N_i} x_i^m x_j^m \mathbf{v}^T \mathbf{f}_{ij}(\mathbf{y}^m) \big) + \lambda_1 \mathbf{u}^T \mathbf{u} + \lambda_2 \mathbf{v}^T \mathbf{v} \Big] \qquad (7)$$

where λ_1 and λ_2 are nonnegative regularizing constants determined by cross-validation. Since the penalized logarithm pseudo-likelihood in (7) is jointly convex with respect to the parameters $\mathbf{u}$ and $\mathbf{v}$, it can be easily minimized using gradient descent.

For a given new test knee MR image $\mathbf{y}'$, we perform the LBP inference algorithm [12] to obtain an optimal label configuration $\mathbf{x}^*$ using the learned model parameters $\{\mathbf{u}, \mathbf{v}\}$. In the implementation, we utilize the UGM toolbox provided by Dr. Mark Schmidt [13].

3. EXPERIMENTS AND RESULTS

In this section, we validate our cartilage segmentation framework with multi-contrast MR images and compare it to SVM [5] and DRF [6]. As in some works on segmenting medical images [8], we employ a subject-specific training scenario, where the classifier is trained on a subset of the subject's data and then tested on another subset or the whole data set. We use every fifth slice as the training data and all the slices as the testing data.

3.1. Evaluation Metrics

Besides visual evaluation, the cartilage segmentations automatically obtained are quantitatively compared to the golden standard using sensitivity, specificity and Dice similarity coefficient (DSC), since they have been used in existing cartilage segmentation works [2, 3, 4]. Sensitivity $= TP / (TP + FN)$ is the true positive fraction and specificity $= TN / (FP + TN)$ the true negative fraction, while $\mathrm{DSC} = 2TP / (2TP + FN + FP)$ is a spatial overlap index, where TP is true positive, TN is true negative, FP is false positive, and FN is false negative counts for the pixels.

3.2. Results

To validate the performance improvements through the use of multi-contrast MR sequences and different types of features, we construct five different feature vectors for each pixel as follows.

- Feature vector 1 (FV1) consists of normalized intensities of multi-contrast MR images only;
- Feature vector 2 (FV2) consists of normalized intensities of multi-contrast MR images and features related to local structure;
- Feature vector 3 (FV3) consists of normalized intensities of multi-contrast MR images and geometrical information based features;
- Feature vector 4 (FV4) consists of normalized intensities of *single* contrast MR images, features related to local structure, and geometrical information based features;
- Feature vector 5 (FV5) consists of normalized intensities of multi-contrast MR images, features related to local structure, and geometrical information based features.

Figure 1 depicts the segmentations of our method for one slice of a knee (It should be noted that we here only illustrate the segmentation results of one knee visually as an example due to the limited space) using FV1, FV2, FV3, FV4 and FV5, respectively, Table 1 gives the average quantitative measures for all slices of all knees. From both visual evaluation (Figure 1) and quantitative evaluation (Table 1), one can see that our method with FV5 gives the best segmentation performance, which indicates the benefit of using multi-contrast MR images and more types of features. Therefore, we will perform the remaining experiments using FV5 only. To compare segmentation performances of DRF, SVM and our method visually and quantitatively, Figure 2 gives the segmentations for one slice of the same knee in Figure 1, the curve of the DSC values for all slices is shown in Figure 3, and Table 2 provides average quantitative measures for all slices. It is easy to observe the superiority of combining DRF with SVM to using only DRF or SVM as the classifier.

4. CONCLUSIONS

In this paper, we utilized DRF to incorporate spatial dependencies in the pixel classification for automatic cartilage segmentation with multi-contrast MR images. We adopted the pixel classification framework based on energy minimization of the convex combination of the SVM-based association potential and the DRF-based interactive potential. We also employed a feature set encoding diverse forms of image and anatomical structure information to improve the segmentation performance. We finally performed the LBP inference algorithm to find the optimal label configuration. Experiments on the human knee multi-contrast MR data show the effectiveness of combining DRF with SVM for cartilage segmentation.

5. ACKNOWLEDGEMENT

We gratefully acknowledge the funding from Ministry of Education, Singapore (Grant No. RG25/08), and also appreciate scholarship support for Kunlei Zhang from China Scholarship Council.

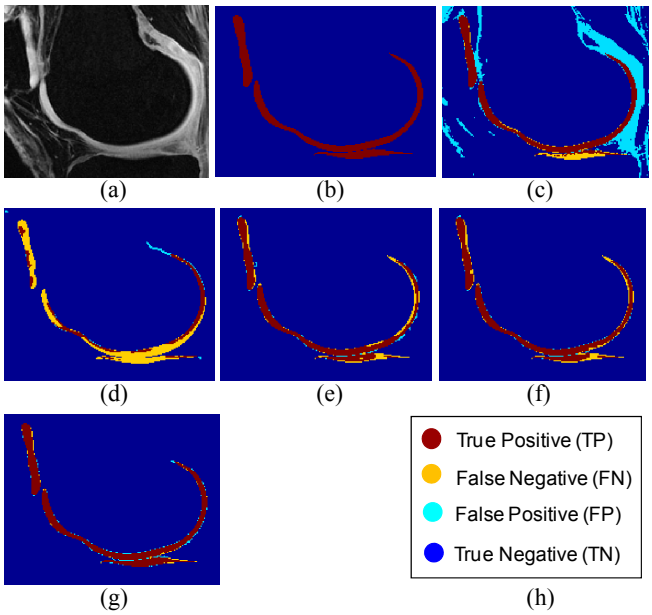

Figure 1. Segmentation results using our method: (a) the cropped MR image for one slice from one knee, (b) the gold standard, (c) segmentation using FV1, (d) segmentation using FV2, (e) segmentation using FV3, (f) segmentation using FV4, (g) segmentation using FV5, (h) the legend.

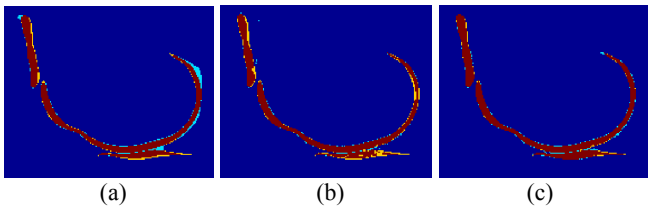

Figure 2. Segmentation results using DRF, SVM and our method: (a) segmentation using DRF, (b) segmentation using SVM, (c) segmentation using our method. Note that the cropped MR image for this slice is the same as in Figure 1(a) and the gold standard is the same as in Figure 1(b).

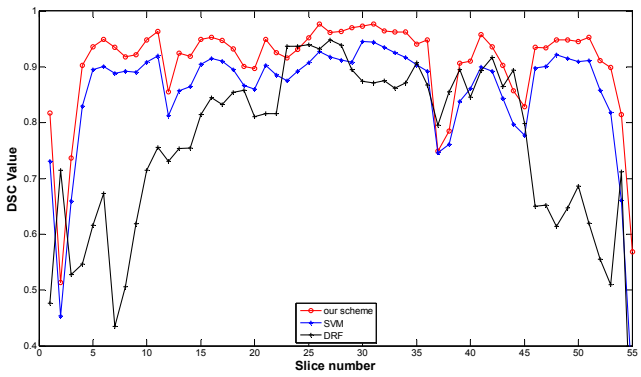

Figure 3. DSC results obtained for all slices from one knee with DRF, SVM and our method, respectively.

Table 1. Mean±std of sensitivity, specificity, and DSC with our method using FV1, FV2, FV3, FV4 and FV5.

our method with	sensitivity	specificity	DSC
FV1	0.843±0.227	0.892±0.008	0.325±0.106
FV2	0.442±0.169	0.995±0.021	0.536±0.238
FV3	0.843±0.129	0.989±0.007	0.866±0.101
FV4	0.879±0.113	0.996±0.002	0.886±0.093
FV5	0.909±0.114	0.997±0.002	0.913±0.090

Table 2. Mean±std of sensitivity, specificity, and DSC using FV5 with DRF, SVM and our method.

method	sensitivity	specificity	DSC
DRFs	0.848±0.095	0.975±0.022	0.775±0.162
SVMs	0.884±0.111	0.996±0.003	0.872±0.112
our method	0.909±0.114	0.997±0.002	0.913±0.090

6. REFERENCES

[1] B. Glocker, N. Komodakis, N. Paragios, C. Glaser, G. Tziritas, and N. Navab, "Primal/dual linear programming and statistical atlases for cartilage segmentation," in *Proc. Int. Conf. MICCAI*, pp. 536-543, 2007.

[2] J. Fripp, S. Crozier, S. K. Warfield, and S. Ourselin, "Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee," *IEEE Trans. Med. Imag.*, vol. 29, pp. 55-64, Jan. 2010.

[3] P. Dodin, J. Pelletier, J. Martel-Pelletier, and F. Abram, "Automatic Human Knee Cartilage Segmentation from 3-D Magnetic Resonance Images," *IEEE Trans. Biomed. Eng.*, vol. 57, pp. 2699-2711, Nov. 2010.

[4] J. Folkesson, E.B. Dam, O.F. Olsen, P.C. Pettersen, and C. Christiansen, "Segmenting articular cartilage automatically using a pixel classification approach," *IEEE Trans. Med. Imag.*, vol. 26, pp. 106-115, Jan. 2007.

[5] S. Koo, P. Alto, B. A. Hargreaves, and G. E. Gold, "Automatic segmentation of articular cartilage from MRI," *Patent*, US 2009/0306496, Dec. 2009.

[6] S. Kumar and M. Hebert, "Discriminative Random Fields," *Int. J. of Comp. Vision*, vol. 68, pp. 179-201, Feb. 2006.

[7] Y. Artan, D. L. Langer, M.A. Haider, T.H. Kwast, A.J. Evans, M.N. Wernick, and I.S. Yetik, "Prostate cancer localization with multispectral MRI using cost-sensitive support vector machines and conditional random fields," *IEEE Trans. Image Processing*, vol. 19, pp. 2444-2455, Sep. 2010.

[8] C.H. Lee, M. Schmidt, A. Murtha, A. Bistritz, J. Sander, and R. Greiner, "Segmenting brain tumor with conditional random fields and support vector machines," in *Proc. Workshop Comp. Vis. Biomed. Image Appl.*, pp. 469-478, Oct. 2005.

[9] http://www.itk.org.

[10] V.N. Vapnik and V. Vapnik, "Statistical Learning Theory," New York: Wiley, 1998.

[11] T. Joachims, "Making large-Scale SVM Learning Practical Advances in Kernel Methods - Support Vector Learning", In *B. Schölkopf and C. Burges and A. Smola, eds: Advances in Kernel Methods – Support Vector Learning*, MIT Press, 1999.

[12] K.P. Murphy, Y. Weiss and M.I. Jordan, "Loopy belief propagation for approximate inference: an empirical study," in *Proc. of Uncertainty in AI*, pp. 467-475, 1999.

[13] http://people.cs.ubc.ca/~schmidtm/Software/UGM.html.