

Higher-Order CRF Tumor Segmentation with Discriminant Manifold Potentials

Samuel Kadoury¹, Nadine Abi-Jaoudeh², and Pablo A. Valdes³

¹ MEDICAL, École Polytechnique de Montréal, Montréal, Canada

² Rad. and Imaging Sciences, National Institutes of Health, Bethesda, MD, USA

³ Dartmouth College, Hanover, New Hampshire, USA

Abstract. The delineation of tumor boundaries in medical images is an essential task for the early detection, diagnosis and follow-up of cancer. However accurate segmentation remains challenging due to presence of noise, inhomogeneity and high appearance variability of malignant tissue. In this paper, we propose an automatic segmentation approach using fully-connected higher-order conditional random fields (HOCRF) where potentials are computed within a discriminant Grassmannian manifold. First, the framework learns within-class and between-class similarity distributions from a training set of images to discover the optimal manifold discrimination between normal and pathological tissues. Second, the conditional optimization scheme computes non-local pairwise as well as pattern-based higher-order potentials from the manifold subspace to recognize regions with similar labelings and incorporate global consistency in the inference process. Our HOCRF framework is applied in the context of metastatic liver tumor segmentation in CT images. Compared to state of the art methods, our method achieves better performance on a group of 30 liver tumors and can deal with highly pathological cases.

1 Introduction

Fast and accurate segmentation of tumors in medical images is an important yet challenging task in many clinical applications, such as radiotherapy or neurosurgery, where follow-up of the tumor size is done with predefined criterions to measure the progression. In various approaches, discrete Markov Random Fields (MRFs; [1]) and Conditional Random Fields (CRFs; [2]) were used to represent complex dependencies among data instances, giving them higher segmentation accuracy than independent and identically distributed (iid) classifiers such as Support Vector Machines (SVM). But despite significant intra- and inter-rater variabilities and the large time consumption of manual segmentation, very few of these automatic approaches are currently used in clinical practice.

State of the art tumor segmentation methods combine efficient classification techniques with low level segmentation methods. From such perspective, tumor detection is addressed as a classification problem where one aims at separating healthy from diseased tissues at the voxel level, while imposing smoothness in the process. In [3], SVM classification using multispectral intensities and

textures is combined with hierarchical CRFs to segment brain tumors in MR images. In [4], a semi-automatic segmentation of liver tumors is presented using hidden Markov measure fields and non-parametric distribution estimation which are defined locally. Alternatively, allowing fully-connected pairwise models can increase expressivity over typical 4 or 8-connected MRF models. Still, the inability to handle higher-order terms is a disadvantage. Higher-order information was shown to enforce labelling consistency over homogeneous regions, as demonstrated with Pn-Potts models [5]. Co-occurrence relations between image classes have also been shown to provide important priors for segmentation [6].

A major drawback from these methods is that conditional probabilities are obtained directly from the high-dimensional image space, which is not ideal as they fail to express the underlying representation of the dataset and assimilates all measures to Euclidean distances [7]. They also assume linear discrimination (SVMs) when in most cases, sets are not linearly separable in the image space. In contrast, manifold learning techniques intrinsically consider the nonlinear distribution of the data, and allow relevant comparison of test cases to the learned population through the use of a mapping distance. Recently, various approaches have used manifolds to track organ motion or discover regional variations within images [8]. However, techniques such as Laplacian eigenmaps are sensitive to outliers and unable to cope with pathological or abnormal tissues. To the best of our knowledge, prior information captured by a discriminant embedding has yet to be exploited in a fully automatic CRF segmentation framework. We propose a segmentation approach for tumors using fully-connected higher-order conditional random field (HOCRF) which maximizes the full potential of high-level graphical models and manifold embeddings. Our contributions are two-fold. First, a discriminant graph-embedding with Grassmannian kernels is generated from prior data to learn the nonlinear intensity distributions of normal and diseased tissue. Second, we suggest to employ a recently proposed mean-field CRF inference approach where potentials are computed directly from the low-dimensional embedding, capturing the underlying structure. Unary and pairwise potentials assess the proximity to manifold regions and the dissimilarity between pairs of segments, respectively. Higher-order potentials ensure regional consistency to efficiently discriminate tumors from healthy tissue. We use the manifold constrained HOCRF segmentation method on metastatic liver tumors in CT images and show the potential of the approach on high-grade gliomas in MRI.

2 Learning the Discriminant Grassmannian Manifold

Manifold learning algorithms are based on the premise that data are often of artificially high dimension and can be embedded in a lower dimensional space. However the presence of outliers and multi-class information can adversely affect the discrimination and/or generalization ability of the manifold. We propose to first learn the optimal separation between normal and pathological tissue by using a discriminant graph-embedding based on Grassmannian manifolds in the segmentation process. Each sample point x_i (pixel or segment) on a Grassmannian manifold can be viewed as the set of m -dimensional subspaces of \mathbb{R}^D and represented

by orthonormal matrices, each with a size of $D \times m$. Two points on a Grassmannian manifold are equivalent if one can be mapped into the other one by a $m \times m$ orthogonal matrix. In this work, similarity between two points (x_i, x_j) on the manifold is measured as a combination of projection and canonical correlation Grassmannian kernels $\mathbb{K}_{i,j}$ defined in the Hilbert Space. With each kernel describing different features of the tumor and normal regions, $\mathbb{K}_{i,j}$ covers a wide spectrum of feature distributions and improves discriminatory accuracy.

In order to effectively discover the low-dimensional embedding, it is necessary to maintain the local structure of the data in the new embedding. The structure $G = (\mathbf{V}, \mathbf{W})$ is an undirected similarity graph, with a collection of nodes \mathbf{V} connected by edges, and the symmetric matrix \mathbf{W} with elements describing the relationships between the nodes. The diagonal matrix \mathbf{D} and the Laplacian matrix \mathbf{L} are defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$, with $\mathbf{D}(i, i) = \sum_{j \neq i} \mathbf{W}_{ij} \forall i$. Here, N labelled points $\mathbb{X} = \{(x_i, c_i)\}_{i=1}^N$ are generated from the underlying manifold \mathcal{M} , where c_i denotes the label (lesion or non-lesion). The task at hand is to maximize a measure of discriminatory power by mapping the underlying data into a vector space, while preserving similarities between data points in the high-dimensional space. Discriminant graph-embedding based on locally linear embedding (LLE) [7] uses graph-preserving criteria to maintain these similarities, which are included in a sparse and symmetric $N \times N$ matrix, denoted as M .

Within and between Similarity Graphs: In our work, the geometrical structure of \mathcal{M} can be modeled by building a within-class similarity graph \mathbf{W}_w for tissues of same type and a between-class similarity graph \mathbf{W}_b , to separate normal and tumor tissue. When constructing the discriminant LLE graph, elements are partitioned into \mathbf{W}_w and \mathbf{W}_b classes. The intrinsic graph G is first created by assigning edges only to vertices of the same class (lesion or non-lesion). The local reconstruction coefficient matrix $M(i, j)$ is obtained by minimizing:

$$\min_M \sum_{j \in \mathcal{N}_w(i)} \|x_i - M(i, j)x_j\|^2 \quad \sum_{j \in \mathcal{N}_w(i)} M(i, j) = 1 \quad \forall i \tag{1}$$

with $\mathcal{N}_w(i)$ as the neighborhood of size k_1 , within the same region as point i (e.g. tumor region). Each sample is therefore reconstructed only from images of the same region. The local reconstruction coefficients are incorporated in the within-class similarity graph, such that the matrix \mathbf{W}_w is defined as:

$$W_w(i, j) = \begin{cases} (M + M^T - M^T M)_{ij}, & \text{if } x_i \in \mathcal{N}_w(x_j) \text{ or } x_j \in \mathcal{N}_w(x_i) \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Conversely, the between-class similarity matrix \mathbf{W}_b depicts the statistical properties to be avoided in the optimization process and used as a high-order constraint. Distances between healthy and pathological samples are computed as:

$$W_b(i, j) = \begin{cases} 1/k_2, & \text{if } x_i \in \mathcal{N}_b(x_j) \text{ or } x_j \in \mathcal{N}_b(x_i) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

with \mathcal{N}_b containing k_2 neighbors having different class labels from the i th sample. The objective is to transform points to a new manifold \mathcal{M} of dimensionality d , i.e. $x_i \rightarrow y_i$, by mapping connected healthy or tumor samples in \mathbf{W}_w as close as possible to the class cluster, while moving tumor and healthy areas of \mathbf{W}_b as far away from one another. This results in optimizing the objective functions:

$$f_1 = \min \frac{1}{2} \sum_{i,j} (y_i - y_j)^2 W_w(i, j) \quad f_2 = \max \frac{1}{2} \sum_{i,j} (y_i - y_j)^2 W_b(i, j) \quad (4)$$

Supervised Manifold Learning: The optimal projection matrix, mapping new points to the manifold, is obtained by simultaneously maximizing class separability and preserving interclass manifold property, as described by the objective functions in Eq.(4). Assuming points on the manifold are known as similarity measures given by the Grassmannian kernel $\mathbb{K}_{i,j}$, a linear solution can be defined, i.e., $y_i = (\langle \alpha_1, x_i \rangle, \dots, \langle \alpha_r, x_i \rangle)^T$ for the r largest eigenvectors with $\alpha_i = \sum_{j=1}^N a_{ij} x_j$. Defining the coefficient $\mathbf{A}_i = (a_{i1}, \dots, a_{iN})^T$ and kernel $\mathbf{K}_i = (k_{i1}, \dots, k_{iN})^T$ vectors, the output can be described as $y_i = \langle \alpha_i, x_i \rangle = \mathbf{A}_i^T \mathbf{K}_i$. By replacing the linear solution in the minimization and maximization of the between- and within-class graphs, the optimal projection matrix \mathbb{A} is acquired from the optimization of the function as proposed by Harandi [9]. The proposed algorithm uses the points on the Grassmannian manifold implicitly (i.e., via measuring similarities through a kernel) to obtain a mapping \mathbb{A} . The matrix maximizes a quotient similar to discriminant analysis, while retaining the overall geometrical structure. Hence for any test point x_q , a manifold representation $v_q = \mathbb{A}^T \mathbf{K}_q$ is obtained using the kernel function based on x_q and mapping \mathbb{A} .

3 Fully-Connected Higher-Order CRF Segmentation

Once the Grassmannian manifolds are obtained from the training phase, the segmentation problem is performed on a higher-order CRF model where potentials of the energy function are inferred from the low-dimensional embeddings. Random variables \mathcal{X} denotes whether labels belong to the object of interest (lesion or not), \mathcal{E} denotes edges connecting pairs of nodes and \mathcal{S} denotes the set of segments. The CRF model defines the energy function based on the sum of unary ψ and pairwise potentials ϕ , as well as higher-order functions θ , minimized by:

$$E(\mathcal{C}|\mathcal{X}) = \sum_{v_i \in \mathcal{X}} \psi(c_i|v_i) + \sum_{(v_i, v_j) \in \mathcal{E}} \phi(c_i, c_j|v_i, v_j) + \sum_{s \in \mathcal{S}} \theta_s(c_s|\mathbf{v}_s) \quad (5)$$

where v_i denotes the manifold projection of data point x_i and c_i is the label (lesion or non-lesion) of node v_i . The nodes in the CRF model corresponds to the manifold embedded segments of x_i . Here, image segments x_i are obtained by a normalized-cut segmentation method, followed by an iterative k-means clustering to create an image parcelled in equal-sized patches [10]. Then, segments are mapped using the projection matrix such that $v_i = \mathbb{A}^T \mathbf{K}_i$. Features are selected

implicitly by the manifold from the low-dimensional space to obtain the best discrimination between the classes. Here, we employ a fully-connected CRF model using a mean-field approximation of the original CRF such that the distribution is composed of a set of independent marginals minimizing KL-divergence.

Manifold-Based Potentials: *Unary potentials* in the CRF model are defined based on the probability of the classifier $\psi(c_i|v_i) = -\log(P(c_i|v_i))$. *Pairwise potentials* are expressed as non-parametric manifold dissimilarities, extending the Gaussian kernel formulation in feature space which uses mean-field approximations of fully-connected CRF models [11]. Instead of forcing Euclidean features to fulfill this task, pairwise potentials are conditioned on the input data such that $\phi(v_i, v_j) = \mu(v_i, v_j) \exp(-d(i, i, \mathcal{X}, \mathcal{M}))$, where μ is the label compatibility function between nodes v_i and v_j . The distance between points i and j , under label l , is the conditional distribution of the label l , with $\exp(-d(i, i, \mathcal{X}, \mathcal{M})) = P(v_j = l|v_i = l, \mathcal{X}, \mathcal{M})$. Assuming the Frobenius distance F in manifold space can offer a non-parametric estimation of the dissimilarity measure, the pairwise potential can be defined by imposing a range σ_f over which valuable information can be inferred when applying a Gaussian window:

$$\phi(c_i, c_j|v_i, v_j) = w \exp\left(\frac{\|v_i - v_j\|_F^2}{2\sigma_f^2}\right) \quad (6)$$

with the parameter w weighting the pairwise relations. Finally, quality sensitive *higher-order potentials* define the label inconsistency in regions (different labels are assigned to a set of neighboring segments) by adding edges between nodes that are not immediate neighbours. We use a strategy where for a given clique \mathbf{v}_s grouping t segments, t different embeddings are generated with \mathbb{A} . We then use the variance of the embedded coordinates of v_s for all t data points, using it as a quality measure $G(s) : s \rightarrow \mathbb{R}$ for all consistent segments in \mathbf{v}_s . The higher-order potential is defined as:

$$\theta_s(\mathbf{v}_s) = \begin{cases} N(\mathbf{v}_s) \frac{1}{Q} \gamma_{\max} & \text{if } N(\mathbf{v}_s) \leq Q \\ \gamma_{\max} & \text{otherwise} \end{cases} \quad (7)$$

where $N(\mathbf{v}_s)$ is the number of segments in \mathbf{v}_s not taking the dominant label, γ_{\max} assigns the cost based on $G(s)$ and Q is the truncation parameter. This potential avoids breaking a *good* segment and penalizes other *bad* segments.

Energy Minimization: The higher-order function is solved with α -expansion and $\alpha\beta$ -swap move making algorithms as proposed by Kholi [12]. In order to determine the optimal moves for the algorithm, higher-order move functions are minimized by transforming the function to quadratic submodular functions which add auxiliary binary variables. The solution corresponds to the energy minima yielding the optimal labeling at each segment, generating tumor and non-tumor segmentation. The optimal labeling of the HOCRf is recovered using Fast-PD based on linear programming.

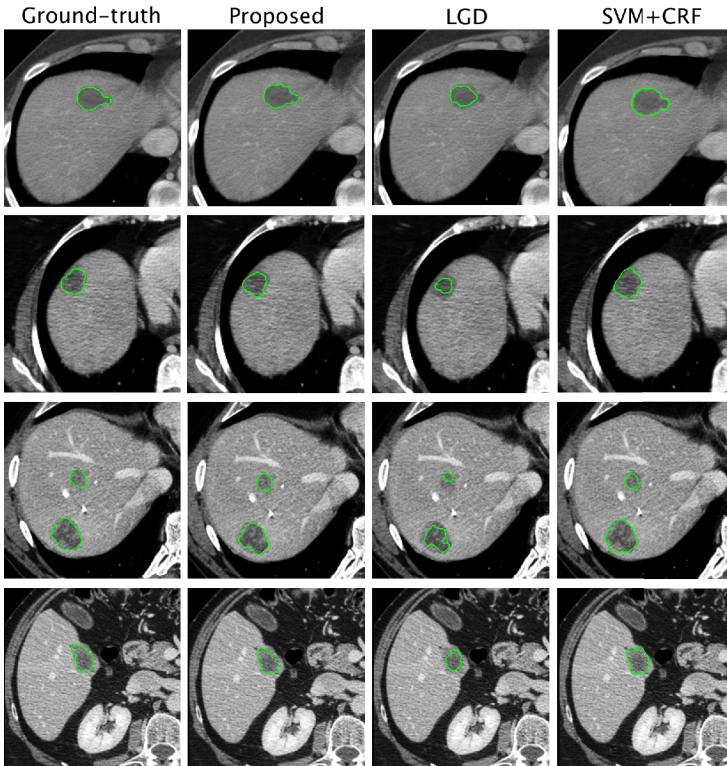


Fig. 1. Segmentation results of metastatic liver tumors in CT images

4 Experimental Validation

We validated our algorithm with the segmentation of metastatic liver tumors in CT images. Datasets from 30 liver tumors were acquired with a 64-slice CT scanner with 1mm slice thickness and in-plane resolution of 0.6mm. Tumors covered only metastasis pathologies. Contrast agent was administered in 12 cases. Tumors were manually segmented by an experimented radiologist and validated by another radiologist. The following measures were computed for each segmentation based on the 2008 liver tumor segmentation challenge: volumetric overlap error (%), average symmetric surface distance (mm), RMS symmetric surface distance (mm) and maximum symmetry surface distance (mm). The Grassmannian manifold was trained with 10 datasets and the optimal dimensionality was found at $d = 5$, when the trend of the nonlinear residual reconstruction error curve stabilized for the entire training set. Table 1 shows the performance of the method with 10 training and 20 test cases. Results were compared to an active contours approach with local Gaussian distributions [13] and to a texture classification method [3]. The average and RMS surface distances (0.6 ± 0.2 and 1.4 ± 0.2 mm) of the proposed method were significantly lower ($p \leq 0.05$) to

Table 1. Error metrics from the CT liver tumor segmentations. We present results using only unary and pairwise ($\psi + \phi$) and unary, pairwise higher-order terms ($\psi + \phi + \theta$).

	Overlap error (%)	Vol. Diff. (%)	Avg. Surf. Dist. (mm)	RMS Surf. Dist. (mm)	Max. Surf. Dist. (mm)
LGD [13]	28.2 ± 3.4	19.6 ± 5.3	1.4 ± 0.3	1.9 ± 0.7	8.1 ± 2.1
SVM+CRF [3]	26.4 ± 2.9	16.0 ± 4.5	1.1 ± 0.3	1.8 ± 0.5	7.8 ± 2.0
HOCRF Training (n=10)	2.1 ± 0.5	0.1 ± 0.1	0.1 ± 0.1	0.4 ± 0.2	1.1 ± 0.9
HOCRF Test $\psi + \phi$ (n=20)	18.5 ± 1.8	12.3 ± 2.9	0.8 ± 0.2	1.5 ± 0.3	6.6 ± 1.7
HOCRF Test $\psi + \phi + \theta$ (n=20)	16.2 ± 1.6	9.8 ± 2.2	0.6 ± 0.2	1.4 ± 0.2	5.9 ± 1.4

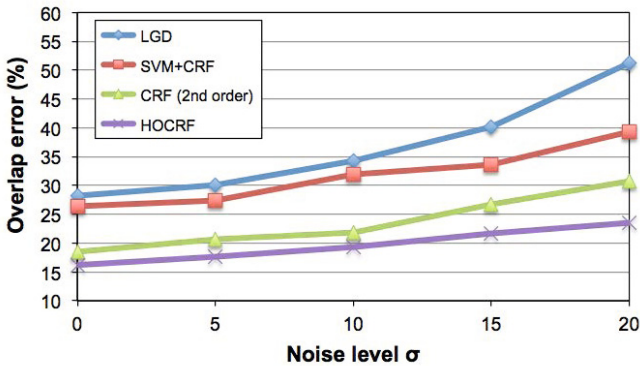


Fig. 2. Evolution of overlap error with increasing Gaussian noise levels added to images

distances generated by [13] and [3]. Typical problems occurred in the periphery of the tumors and in cases of rim-enhancing liver metastases. These cases offer a density which was not observed in the training set, but could be compensated with additional data in the manifold. Fig. 1 shows some segmentation results.

By assessing the gain in accuracy when adding the higher-order terms in the energy formulation, the overlap error is reduced by 2.3%, which is a statistically significant difference ($p \leq 0.05$) to the second-order MRF model. In order to evaluate the robustness of the method, we performed additional experiments by measuring segmentation accuracy with 4 different levels of Gaussian noise added to the input images. Fig. 2 demonstrates that the proposed methodology possesses increased tolerance to noise compared to the other methods.

5 Conclusion

We proposed a new, efficient and adaptable method for tumor segmentation. This was achieved through a higher-order fully-connected graphical model that was optimized using potential functions defined in a discriminant Grassmannian manifold. This increased the ability to isolate diseased from healthy tissue as compared to state of the art CRF or SVM techniques. Validation concerning

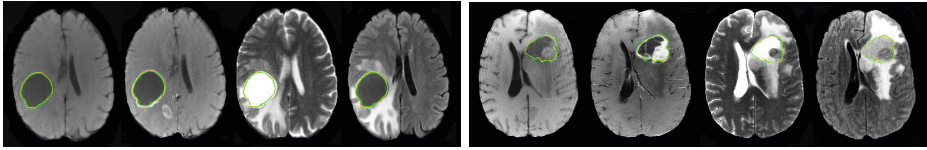


Fig. 3. Sample high-grade tumor segmentations in brain MRI for 2 glioma cases

the case of metastatic liver tumors was considered to evaluate the performance of the method, leading to very promising results. The approach can also be extended to other pathologies such as for example glioblastoma multiforme by segmenting high-grade gliomas in MRI. In this case, the manifold potentials could be trained to characterize different necrotic, active and edema regions in the tumor, and help to automatically discriminate between normal and diseased brain tissue. Preliminary segmentation results for high-grade glioma cases are shown in Fig. 3. Encoding prior knowledge relating to shape representation is a natural extension of the proposed formulation. Future work will focus on a more extensive validation of the segmentation process for gliomas in brain MRI.

References

1. Li, S.Z.: Markov Random Field Modeling in Image Analysis. Springer (2009)
2. Lafferty, J., Pereira, F., McCallum, A.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: ICML, pp. 282–289 (2001)
3. Bauer, S., Nolte, L.-P., Reyes, M.: Fully Automatic Segmentation of Brain Tumor Images Using Support Vector Machine Classification in Combination with Hierarchical Conditional Random Field Regularization. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 354–361. Springer, Heidelberg (2011)
4. Hame, Y., Pollari, M.: Semi-automatic liver tumor segmentation with HMM field model and non-parametric distribution estimation. *Med. I. Ana.* 16, 140–149 (2012)
5. Sabuncu, M., Yeo, B., Leemput, K., et al.: A Generative Model for Image Segmentation Based on Label Fusion. *IEEE Trans. Med. Imag.* 29, 1714–1728 (2010)
6. Vineet, V., Warrell, J., Torr, P.H.S.: Filter-based mean-field inference for random fields with higher-order terms and product label-spaces. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part V. LNCS, vol. 7576, pp. 31–44. Springer, Heidelberg (2012)
7. Roweis, S., Saul, L.: Nonlinear dimensionality reduction by locally linear embedding. *Science* 290, 2323–2326 (2000)
8. Bhatia, K.K., Rao, A., Price, A.N., Wolz, R., Hajnal, J., Rueckert, D.: Hierarchical manifold learning. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 512–519. Springer, Heidelberg (2012)
9. Harandi, M., Sanderson, C., et al.: Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching. In: CVPR, p. 2705 (2011)
10. Mori, G.: Guiding model search using segmentation. In: ICCV, pp. 1417–1423 (2005)
11. Krähenbühl, P., Koltun, V.: Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. In: NIPS, vol. 24, pp. 109–117 (2011)
12. Kohli, P., Ladicky, L., Torr, P.: Robust higher order potentials for enforcing label consistency. *IJCV* 82, 302–324 (2009)
13. Wang, L., He, L., Mishra, A., Li, C.: Active contours driven by local gaussian distribution fitting energy. *Signal Process.* 89, 2435–2447 (2009)