Context Enhanced Graphical Model for Object Localization in Medical Images

Yang Song[1], Weidong Cai[1], Heng Huang[2], Yue Wang[3], and David Dagan Feng[1]

[1]BMIT Research Group, School of IT, University of Sydney, Australia
[2]Computer Science and Engineering, University of Texas at Arlington
[3]Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University

Abstract. Object localization is an important step common to many different medical applications. In this Chapter, we will review the challenges and recent approaches tackling this problem, and focus on the work by Song et.al. [20]. In [20], a new graphical model with additional contrast and interest-region potentials is designed, encoding the higher-order contextual information between regions, on the global and structural levels. A discriminative sparse-coding based interest-region detector is also integrated as one of the context prior in the graphical model. This object localization method is generally applicable to different medical imaging applications, in which the objects can be distinguished from the background mainly based on feature differences. Successful applications on two different medical imaging applications – lesion dissimilarity on thoracic PET-CT images and cell segmentation on microscopic images – are demonstrated in the experimental results.

1 Introduction

A wide variety of medical applications comprise object localization as an important step for discovering the anatomical or pathological information from images. For example, region-of-interest (ROI) detection is helpful for early screening of diseases; and lesion segmentation is useful for treatment planning. We consider object localization as a generalization of both detection and segmentation, with both automatic identification of ROI, and a good delineation of its boundary.

We focus on medical imaging problems in which objects can be localized based on local-level features and feature differences between the objects and background. For example, in positron emission tomography – computed tomography (PET-CT) images, abnormalities typically show higher uptakes than normal tissues. In fluorescence microscopic images, the cell nuclei normally depict darker colors then the other cell structures and the background. In brain magnetic resonance imaging (MRI), the white and gray matter display quite different intensities and spatial patterns.

Local features are usually not sufficient for a good localization, because of large inter-subject variations causing same anatomical structures appearing quite

differently across images. The problem is further complicated due to low feature differences between different tissue types and especially for the boundary areas between the objects and background. In addition, pathologies often lead to larger imaging variations, and an accurate object localization is thus more challenging.

For such imaging problems, while lots of work have been reported [25, 16, 19, 5, 18, 4, 21, 15], they are mostly designed to be domain specific; and often rely on sophisticated feature sets, which can be computational-intensive and difficult to adapt to other imaging problems. Furthermore, because such features are usually designed based on domain knowledge and empirical studies, their effectiveness may be restricted to the limited scenarios available in the datasets.

Therefore, in [20], we proposed an object localization method that can be generally applicable, requires simpler feature sets, and addresses low feature differences and large inter-subject variations. With region-based labeling, each image region is classified as the object or background to produce the localization output. In summary, our main contributions are the following: (i) the discriminative capability of the basic conditional random field (CRF) is enhanced with two contextual priors, namely the contrast and interest-region potentials, to encode the global contrast information and region-based feature similarities, for improving the boundary delineations; (ii) a sparse-coding classification method is proposed for interest-region detection, with improved discriminative power of the learned dictionaries; and (iii) the design is kept general with simple feature sets configurable for the specific application, and has been successfully applied to both lesion dissimilarity on thoracic PET-CT images and cell segmentation on microscopic images.

Related Work. We focus our review on CRF-based localization methods in both medical and general imaging domains. As an undirected graphical model, CRF is now one of the most successful trends in object class image segmentation [6]. The basic and most commonly used formulation is to have local features represented as graph nodes and consistency constraints between neighboring regions as edge connections [17]. However, comparing to the non-graphical discriminative approach, generally such models add advantages little more than spatial smoothing of labelings [25].

Higher-level features, i.e. contexts in images, are often acknowledged as important discriminative factors [6, 4]. In particular, relationship information on a larger scale, such as those across image slices [8], relating to reference objects [2], or between distant image regions [7], can be modeled as pairwise connections to encourage labeling consistency or enhance the discriminative power of local features. Such ideas of connecting beyond immediate neighbors are inspiring; however, choosing the related pairs and describing their interactions are rather application specific. To explore multi-scale region interactions, hierarchical models have been proposed [11, 3]; however, they are normally created based on region clustering, without considering the actual object structures. At a more structural level, object detectors with bounding box outputs have been incorporated into CRFs as consistency constraints [12, 6]. Although the idea is sound,

such methods are normally built based on well-established object detectors and thus require only simple interaction modeling; but both assumptions are not suitable for our problem domain.

2 Object Localization

Given an image I, we first oversegment it into a set of regions $\{r_p\}$, using quick-shift clustering [24], to incorporate superpixel-level information around the pixels. The objective of object localization is then to derive a binary mask $L = \{l_p\}$, with each $l_p \in \{0, 1\}$ indicating whether the region r_p belongs to the object.

2.1 The Proposed CRF Model

We formulate the object localization problem as a binary labeling task using a new CRF model, with the following energy function:

$$E(L|I) = \underbrace{\sum_{p} \phi_L(l_p)}_{\text{local}} + \underbrace{\sum_{(p,q)\in N_S} \psi_S(l_p, l_q)}_{\text{smooth}} + \underbrace{\sum_{(p,c)\in N_C} \psi_C(l_p, l_c)}_{\text{contrast}} + \underbrace{\sum_{(p,i)\in N_R} \psi_R(l_p, l_i)}_{\text{interest-region}} \tag{1}$$

where the set of random variables or nodes of the graph is denoted by $L = \{\{l_p\} \cup \{l_c\} \cup \{l_i\}\}$, including the new auxiliary nodes from the contrast (l_c) and interest-region (l_i) potentials. The probability of a certain configuration is a conditional distribution on the energy function $E(L|I)$, and the optimal labeling is derived by minimizing the total energy using the graph cut [10].

The local potential $\phi_L(l_p)$ describes the cost of r_p labeled as 0 or 1:

$$\phi_L(l_p) = 1 - P(r_p = l_p|f_p) \tag{2}$$

where f_p is the local feature vector of r_p, and $P(.)$ is the probability estimate of labeling obtained using a binary support vector machine (SVM).

The smooth potential $\psi_S(r_p, r_q)$ penalizes the differences in labeling of the neighboring regions r_p and r_q based on their feature distances with a Potts model:

$$\psi_S(l_p, l_q) = \exp(-\frac{\parallel f_p - f_q \parallel^2}{2\beta_S})\mathbf{1}(l_p \neq l_q) \tag{3}$$

where β_S is the normalization factor as the average of all L2 distances between neighboring feature vectors in I. The regions r_p and r_q are considered neighbors if they share some common border in I, and the set of all neighboring pairs is denoted by N_S.

While the first two potentials follow the standard CRF constructs (Figure 1a), we describe the contrast and interest-region potentials (ψ_C, ψ_R, N_C and N_R) in the following.

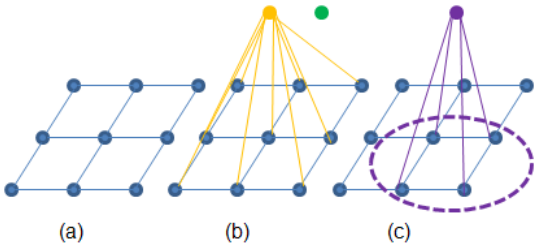


Fig. 1. The proposed CRF model. (a) The standard CRF construct, with nodes representing the image regions and edges linking the neighboring regions. (b) Introducing two auxiliary nodes (object and background) for the contrast potential, with edges linking the image regions and the auxiliary nodes (showing only one set of edges for easier viewing). (c) Based on the detected interest region (purple circle), an auxiliary node for the interest-region potential is added, with edges linking all image regions in the interest-region and the added node.

2.2 Contrast Potential

To improve the labeling accuracy, we want to explore the contrast information in the image I, with the following motivations. Across different images, there are often large inter-subject variations, causing overlaps between the feature ranges and hence misclassifications. Nevertheless, within one image, there is always a certain degree of contrast between the objects and background; and the contrast information helps to discriminate between the two types. To encode the contrast information, two additional nodes corresponding to the object and background, namely the contrast nodes l_c^o and l_c^b, are then added to the graph. A pairwise connection between the image region l_p and each of the two nodes is also established (Figure 1b), and N_C denotes the set of all such pairwise connections. With such a construct, we expect to encourage the same labelings between the image region and contrast nodes if they exhibit similar features, and also different labelings otherwise.

To do this, we first define the unary potentials of the two contrast nodes:

$$\phi_C(l_c^{o/b}) = \left\{ \begin{array}{ll} 0 & \text{if } l_c^{o/b} = 1/0 \\ C & \text{otherwise} \end{array} \right\} \tag{4}$$

where C is a large constant, so that large costs are assigned to $l_c^o \neq 1$ and $l_c^b \neq 0$ and 0 costs otherwise, to effectively fix the labelings of the two nodes in the inference results.

We then define the pairwise potentials for the edges (l_p, l_c) with the following. First, based on the labeling outputs with local features only (Eq. (2)), we obtain the initial estimation of the objects and background areas, and two feature vectors f_c^o and f_c^b are then derived for the estimated objects and background (details of feature derivation in Sec 4). Next, we compute the contrast features

between r_p and the objects and background as $g_p = \{f_p, f_p/f_c^o, f_p/f_c^b\}$, and classify the feature g_p to two classes – likely or unlikely to represent the object, denoted as $likely(1)$ and $unlikely(1)$ – using a binary SVM. Then, based on the probability estimates γ_p of class $likely(1)$, the pairwise costs are computed as:

$$\psi_C(l_p, l_c^o) = \left\{ \begin{array}{ll} 0 & \text{if } l_p = 1, \text{ and } likely(1) \\ 1-\gamma_p & \text{if } l_p = 1, \text{ and } unlikely(1) \\ \gamma_p & \text{if } l_p = 0 \end{array} \right\} \tag{5}$$

$$\psi_C(l_p, l_c^b) = \left\{ \begin{array}{ll} 0 & \text{if } l_p = 0, \text{ and } unlikely(1) \\ \gamma_p & \text{if } l_p = 0, \text{ and } likely(1) \\ 1-\gamma_p & \text{if } l_p = 1 \end{array} \right\} \tag{6}$$

Note that because of the *likely* and *unlikely* terms, the above pairwise potentials no longer follow the Potts model, and penalize labeling consistency if the features of the image regions and the contrast nodes are actually dissimilar. The total energy of the contrast potential can however, be rewritten in the following format, to keep it submodular (binary and with pairwise term encouraging consistency) for efficient graph-cut energy minimization:

$$\begin{array}{l} \sum_{(p,c)\in N_C} \psi_C(l_p, l_c) = \sum_c \phi_C(l_c) + \\ \sum_p \alpha_p \mathbf{1}(unlikely(l_p)) + \sum_{(p,c)\in N_C} \alpha_p \mathbf{1}(l_p \neq l_c) \end{array} \tag{7}$$

where $\alpha_p = \gamma_p$ if $l_p = 0$, and $\alpha_p = 1 - \gamma_p$ otherwise.

2.3 Interest Region Potential

Although the contrast nodes represent the object and background regions of an image I on a global scale, the structural information between image regions is not explored. An obviously important structural information is that, regions that are likely parts of the same anatomical or pathological structure should take the same labelings. In our formulation, the hypothesis is that, if we can detect a set of structures, i.e. interest regions R_i, the comprising regions $r_p \in R_i$ should preferably be assigned to the same category, but also depending on their individual suitability of such an labeling. The advantage of such an approach is that, we can employ a totally different method to detect the interest regions (e.g. non-CRF and different features), so the generated regions can serve as a second opinion to refine the object localization.

Assume a set of interest regions R_i are detected from an image I (details in Sec 3), and each interest region is characterized by its feature f_i, most probable label $l_i^* \in \{0, 1\}$ and a set of image regions r_p covered. Note that r_p might partially overlap with R_i especially around the boundary areas of R_i, and hence not all r_p covered by R_i should have the same label as l_i^*. To determine the the probability of $l_p = l_i^*$, we first compute the following feature vector:

$$v_p = \{\cap(r_p, R_i)/r_p, \parallel f_p - f_i \parallel, f_{i-p}/f_i\} \tag{8}$$

which represents the degrees of area overlap and feature homogeneity between r_p and R_i, with f_{i-p} denoting the feature of R_i excluding r_p. Then a binary SVM is trained to classify v_p into *same* or $diff$ categories, indicating if $l_p = l_i^*$ or otherwise, and the probability estimate of $l_p = l_i^*$ is denoted by $\theta_{p,i}$.

Next, to integrate the interest-region detection hypothesis into the CRF formulation, for each R_i detected, a node l_i is added to the graph, with the unary potential $\phi_R(l_i)$ defined similarly to Eq. (4). An edge is then connected between each pair of (l_p, l_i) for all $r_p \in R_i$ (Figure 1c) with N_R denoting all such edges for image I, and we define the pairwise potential as:

$$\psi_R(l_p, l_i) = \theta_{p,i}\mathbf{1}(l_p \neq l_i) \tag{9}$$

Since $r_p \in R_i$ is quite likely to exhibit the same labeling as R_i, we choose to use the Potts model to encourage such consistency. The cost of different labelings is directly related to the probability of $l_p = l_i^*$, and hence we use $\theta_{p,i}$ as the pairwise cost. If r_p is less likely to be labeled as l_i^*, the use of $\theta_{p,i}$ is also able to lessen the consistency constraint.

With the above definitions, the total energy term of the interest-region potential is thus rewritten as the following:

$$\sum_{(p,i)\in N_R} \psi_R(l_p, l_i) = \sum_i \phi_R(l_i) + \sum_{(p,i)\in N_R} \theta_{p,i}\mathbf{1}(l_p \neq l_i) \tag{10}$$

2.4 Graph Inference

All energy terms are given equal weights (based on our empirical evaluation), and piecewise learnings of the probability estimates used in the local, contrast and interest-region potentials are conducted first. The binary inference problem $L^* = \text{argmin } E(L|I)$ is then solved efficiently using the graph cut.

3 Detection for Interest Region Potential

Due to our motivation of detecting the interest regions in a totally different way from the graph-based approach to support the interest-region potential (Sec 2.3), we choose to design a sparse-coding based classification method for interest-region detection. Besides its popularity and widely demonstrated effectiveness [14], we believe sparse coding can be particularly suitable for our problem because of the large variations in the dataset.

3.1 Sparse Coding for Classification

Let Y be a set of n-dimensional data samples $Y = \{y_j : j = 1, ..., J\}$ and $Y \in R^{n\times J}$. A representative dictionary for Y with K atoms is denoted as $D = \{d_k : k = 1, ..., K\} \in R^{n\times K}$. Each y_j can then be represented as a linear combination of a few (i.e. $\leq T$) atoms in D with minimum reconstruction error, and the

corresponding coefficient vector x_j is the sparse code. Denoting the set of sparse codes of the data samples Y as $X = \{x_j : j = 1, ..., J\} \in R^{K \times J}$, both the dictionary D and sparse coding X can be learned with K-SVD [1] by solving the following problem:

$$\langle D, X \rangle = \underset{D,X}{\operatorname{argmin}} \ \|Y - DX\|_2^2 \quad s.t. \forall j, \|x_j\|_0 \leq T \tag{11}$$

where $\|Y - DX\|_2^2$ represents the reconstruction error.

Once the dictionary D is learned, a given data sample y can then be represented as a sparse code x by solving the following using the OMP algorithm [23]:

$$x = \underset{x}{\operatorname{argmin}} \ \|y - Dx\|_2^2 \quad s.t. \ \|x\|_0 \leq T \tag{12}$$

A classifier (e.g. SVM) can then be trained based on a set of such sparse codes, so that x and hence y can be classified.

In our context, an image I is divided into grid-based patches, and each image patch is represented by its feature descriptor y. The dictionary D is generated with a training set Y, and each image patch is then classified as interest region or otherwise ($h \in \{1, 0\}$) based on its sparse code x.

3.2 Discriminative Sparse Learning

A shortcoming with the above approach is the separation of the dictionary learning and classifier training, hence the learned dictionary might not produce discriminative sparse codes for the classification. Several approaches have thus been proposed to integrate the two steps of learning [9]. However, it is observed that such an integrated approach is still largely optimized for the reconstruction term, which may affect the discriminative power of W. Therefore, we suggest that the integrated learning for dictionary D should not totally replace the separate classifier training, and propose a different method as follows.

First, for the data samples $Y \in R^{n \times J}$, we create a corresponding labeling vector $H = \{h_j\} \in \{-1, 1\}^{1 \times J}$, with 1 for interest region. Based on linear-kernel SVM, the optimization objective of the weight vector $w \in R^{1 \times K}$ is:

$$\begin{aligned} &\operatorname{argmin}_{w,\xi,b} \ \tfrac{1}{2}\|w\|^2 + C \textstyle\sum_j \xi_j \\ &\qquad s.t. \ \forall j, h_j(w * x_j + b) \geq 1 - \xi_j, \ \xi_j \geq 0 \end{aligned} \tag{13}$$

Combining Eq. (11) and (13), and by simplifying the complexities caused by the inequality constraints on ξ_j and the signed h_j, we relax the formulation based on least squares SVM (LS-SVM) [22] as:

$$\begin{aligned} &\langle D, X, w \rangle = \operatorname{argmin}_{D,X,w} \ \|Y - DX\|_2^2 + \|w\|^2 + \textstyle\sum_j \xi_j^2 \\ &s.t. \ \forall j, \|x_j\|_0 \leq T, \ h_j(w * x_j + b) = 1 - \xi_j \end{aligned} \tag{14}$$

By combining w and b, and substituting ξ_j, the problem is then equivalent to the following:

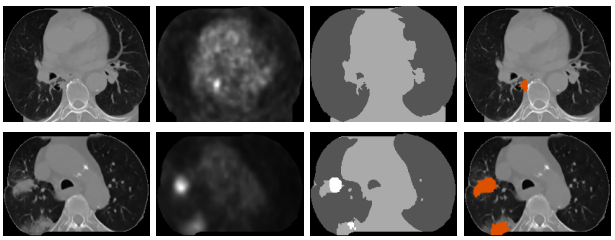
$$\begin{aligned} &\langle D', X', w' \rangle = \text{argmin}_{D',X',w'} \ \|Y - D'X'\|_2^2 + \|w'\|^2 + \\ &\|H - w'X'\|_2^2 \qquad s.t.\ \forall j, \|x'_j\|_0 \leq T \end{aligned} \tag{15}$$

where $w' = [w, b] \in R^{1\times(K+1)}$ and $X' \in R^{(K+1)\times J}$ appended an addition dimension with constant value 1 to absorb b, and $D' \in R^{n\times(K+1)}$ with an additional atom to be dimensionally compatible with X'. To solve Eq. (15), an alternative approach is used, as detailed in [20].

4 Experimental Results

4.1 Results on Lesion Dissimilarity

Measuring lesion similarity is important in many medical applications, such as content-based image retrieval for diagnosis referencing. In our approach, first, lesions (i.e. lung tumors and abnormal lymph nodes) in thoracic PET-CT images are localized in each image slice with the proposed method. Second, their textural and spatial features are extracted in 3D. Lastly, a weighted histogram-intersection is used to compute the feature distance. The actual implementation details are referred to [20]. The datasets comprise of 40 thoracic PET-CT 3D image sets from non-small cell lung cancer studies. A total of 64 lesions including lung tumors and abnormal lymph nodes are annotated, and the similarity/dissimilarity relationships between each pair of 3D image sets are marked as the ground truth. Three image sets showing typical thoracic characteristics are selected for training, and testing is performed on all image sets.



(a) (b) (c) (d)

Fig. 2. Two example localization outputs. (a) Transaxial CT image slices (showing the thorax after preprocessing). (b) Co-registered PET image slices. (c) The labeling outputs using standard CRF, with dark gray for lung field, light gray for mediastinum and white for lesion. (d) Our localization outputs with the two additional potentials, with lesions highlighted in orange.

Figure 2 shows examples of the lesion localization. The first example illustrates the benefits of the contrast potential, in which the lesion is initially not detected with standard CRF, due to the relatively low PET intensities. The interest-region potential is particularly useful in refining the lesion boundaries, which tend to be underestimated with the standard CRF, as shown in the second example. It is observed that, the standard CRF tends to produce a large number of either totally undetected or underestimated lesions. Based on the measured 3D object-level localization results, we summarize the localization recall, precision and F-score in Table 1.

The localized lesions are then used to retrieval images with similar lesions. The retrieval tests are performed by using each 3D image set as a query image, and the remaining 39 images are ranked accordingly. We compare the retrieval performance with three other approaches: (i) state-of-the-art of thoracic PET-CT image retrieval [18]; (ii) spatial pyramid matching (SPM) with local intensity features extracted from grid-based image patches; and (iii) bag-of-words with SIFT descriptor. As shown in Figure 3, our proposed method exhibits the highest retrieval precisions for all recall levels.

Table 1. The localization performances comparing our proposed method with standard CRF.

	Recall (%)	Precision (%)	F-score (%)
Ours	97.0	95.4	96.2
CRF	76.6	94.2	84.5

4.2 Results on Cell Segmentation

Cell nucleus segmentation is one of the most important tasks in analyzing and quantifying fluorescence microscopic images. In our approach, the cell nucleus is localized using the proposed method; and since the localization results also tend to delineate the nucleus boundaries closely, such an approach can be directly used for segmentation. The actual implementation details are referred to [20]. The serous database [13] is used to evaluate the cell segmentation. The database contains 10 microscopic images. A total of 254 cell nuclei are present in the images, with ground truth of cell nuclei segmentation provided. Same as [4], half of the images are used for training and the others for testing.

To evaluate the segmentation performance, we compute the PASCAL VOC criteria of pixel- and object-level accuracies, both as TP/(TP+FN+FP). We also compare our results with three approaches: (i) L+S, the standard CRF with local and smooth potentials; (ii) L+S+C, with additional contrast potential; (iii) L+S+R, with additional interest-region potential; and (iv) the state-of-the-art discriminative labeling method [4] reported for the same database. As listed

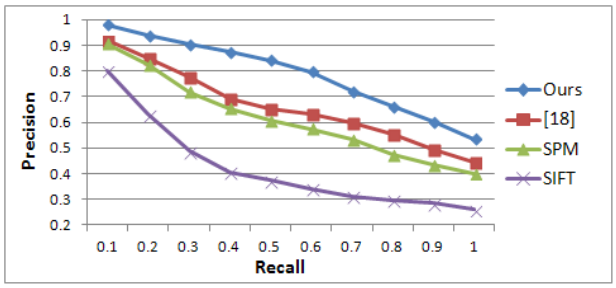


Fig. 3. The retrieval precision and recall.

Table 2. The segmentation results comparing various methods.

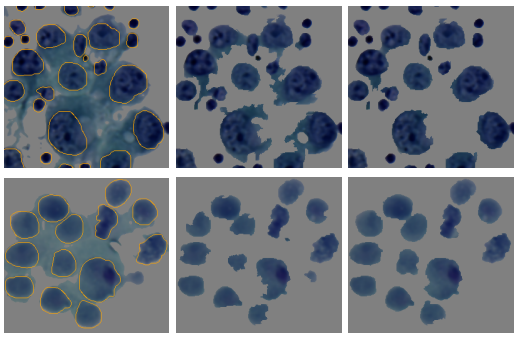
	Ours	L+S	L+S+C	L+S+R	[4]
Pixel Acc (%)	85.6	82.0	83.1	84.6	85.1
Obj Acc (%)	89.3	84.5	86.2	88.7	84.0

in Table 2, our method achieves the highest pixel- and object-level accuracies. The improvements of having the contrast and potential terms are evident. The performance difference between L+S and [4] suggests that if we incorporate the feature set of [4], the segmentation accuracies would be further improved. By replacing the interest-region detection with standard sparse-coding classification, it is found that our proposed method exhibits on average 1.1% improvement for both pixel- and object-level measurements with the new approach.

The first example shown in Figure 4 indicates that our method is quite effective in removing the cytoplasm areas that connect the cell nuclei. As shown in the second example, lighter intensities of the cell nuclei cause many false negatives with the standard CRF approach; and our result shows more accurate delineations of the actual contours.

5 Summary

In this Chapter, we describe a new method for object localization in medical images [20]. A new CRF model with additional contrast and interest-region potentials is proposed for effective object localization, addressing large inter-subject variations and low feature differences between the objects and background. A new sparse-coding classification approach is also designed for the interest-region detection, with enhanced discriminative power of the learned dictionaries. The proposed method is applied to lesion dissimilarity on thoracic PET-CT images, and cell segmentation on microscopic images, and shows higher performance compared to the state-of-the-art techniques.



(a) (b) (c)

Fig. 4. Two example segmentation results. (a) Cropped microscopic images, with orange circles delineating the segmentation ground truth. (b) The segmentation results with L+S. (c) The segmentation results of our proposed method.

References

1. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process. 54(1), 4311–4322 (2006)
2. Ayed, I.B., Punithakumar, K., Garvin, G., Romano, W., Li, S.: Graph cuts with invariant object-interaction priors: application to intervertebral disc segmentation. In: Szekely, G., Hahn, H. K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 221–232. Springer, Heidelberg (2011)
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. Cheng, L., Ye, N., Yu, W., Cheah, A.: Discriminative segmentation of microscopic cellular images. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 637–644. Springer, Heidelberg (2011)
5. Feuerstein, M., Glocker, B., Kitasaka, T., Nakamura, Y., Iwano, S., Mori, K.: Mediastinal atlas creation from 3-d chest computed tomography images: application to automated detection and station mapping of lymph nodes. Med. Image Anal. 16(1), 63–74 (2011)
6. Gonfaus, J., Boix, X.: Harmony potentials for joint classification and segmentation. In: CVPR, pp. 3280–3287 (2010)
7. Guo, R., Dai, Q., Hoiem, D.: Single-image shadow detection and removal using paired regions. In: CVPR, pp. 2033–2040 (2011)
8. Jagadeesh, V., Vu, N., Manjunath, B.S.: Multiple structure tracing in 3D electron micrographs. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 613–620. Springer, Heidelberg (2011)
9. Jiang, Z., Lin, Z., Davis, L.: Learning a discriminative dictionary for sparse coding via label consistent K-SVD. In: CVPR, pp. 1697–1704 (2011)

10. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Trans. Pattern Anal. Mach. Intell. 26(2), 147–159 (2004)
11. Ladicky, L., Russell, C., Kohli, P., Torr, P.H.S.: Associative hierarchical CRFs for object class image segmentation. In: ICCV, pp. 739–746 (2009)
12. Ladicky, L., Sturgess, P., Alahari, K., Russell, C., Torr, P.H.S.: What, where and how many? Combining object detectors and CRFs. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part IV. LNCS, vol. 6314, pp. 424–437. Springer, Heidelberg (2010)
13. Lezoray, O., Cardot, H.: Cooperation of color pixel classification schemes and color watershed: a study for microscopical images. IEEE Trans. Image Process. 11(7), 783–789 (2002)
14. Liu, M., Lu, L., Ye, X., Yu, S., Salganicoff, M.: Sparse classification for computer aided diagnosis using learned dictionaries. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 41–48. Springer, Heidelberg (2011)
15. Lu, C., Chelikani, S., Jaffray, D.A., Milosevic, M.F., Staib, L.H., Juncan, J.S.: Simultaneous nonrigid registration, segmentation, and tumor detection in MRI guided cervical cancer radiation therapy. IEEE Trans. Med. Imag. 31(6), 1213–1227 (2012)
16. van Ravesteijin, V.F., van Wijk, C., Vos, F.M., Truyen, R., Peters, J.F., Stoker, J., van Vliet, L.J.: Computer-aided detection of polyps in CT colonography using logistic regression. IEEE Trans. Med. Imag. 29(1), 120–131 (2010)
17. Shotton, J., Winn, J., Rother, C., Criminisi, A.: Textonboost: joint appearance, shape and context modeling for multi-class object recognition and segmentation. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 1–15. Springer, Heidelberg (2006)
18. Song, Y., Cai, W., Eberl, S., Fulham, M., Feng, D.: Discriminative pathological context detection in thoracic images based on multi-level inference. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 191–198. Springer, Heidelberg (2011)
19. Song, Y., Cai, W., Eberl, S., Fulham, M., Feng, D.: Thoracic image case retrieval with spatial and contextual information. In: ISBI pp. 1885–1888 (2011)
20. Song, Y., Cai, W., Huang, H., Wang, Y., Feng, D.D.: Object localization in medical images based on graphical model with contrast and interest-region terms. In: CVPR Workshop, pp. 1–7 (2012)
21. Song, Y., Cai, W., Kim, J., Feng, D.D.: A multistage discriminative model for tumor and lymph node detection in thoracic images. IEEE Trans. Med. Imag. 31(5), 1061–1075 (2012)
22. Suykens, J., Vandewalle, J.: Least squares support vector machine classifiers. Neural Process. Letters 9(3), 293–300 (1999)
23. Tropp, J.: Greed is good: algorithmic results for sparse approximation. IEEE Trans. Inf. Theory 50(10), 2231–2242 (2004)
24. Vedaldi, A., Soatto, S.: Quick shift and kernel methods for mode seeking. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 705–718. Springer, Heidelberg (2008)
25. Wu, D., Lu, L., Bi, J., Shinagawa, Y., Boyer, K., Krishnan, A., Salganicoff, M.: Stratified learning of local anatomical context for lung nodules in CT images. In: CVPR, pp. 2791–2798 (2010)