Lymph node detection in 3-D chest CT using a spatial prior probability
— **Source link** ↗

Johannes Feulner, S. Kevin Zhou, Martin Huber, Joachim Hornegger ...+2 more authors

Institutions: University of Erlangen-Nuremberg, Princeton University, Siemens

Related papers:

- Automatic detection and segmentation of axillary lymph nodes

- Automated extraction of lymph nodes from 3-D abdominal CT images using 3-D minimum directional difference filter

- Automatic Detection and Segmentation of Lymph Nodes From CT Data

- Hierarchical Parsing and Semantic Navigation of Full Body CT Data

- Mediastinal lymph node detection on thoracic CT scans using spatial prior from multi-atlas label fusion

Lymph Node Detection in 3-D Chest CT using a Spatial Prior Probability

Johannes Feulner[1,3], S. Kevin Zhou[2], Martin Huber[3], Joachim Hornegger[1],
Dorin Comaniciu[2], Alexander Cavallaro[4]

[1]Pattern Recognition Lab, University of Erlangen-Nuremberg, Germany
[2]Siemens Corporate Research, Princeton, NJ, USA
[3]Siemens Corporate Technology, Erlangen, Germany
[4]Radiology Institute, University Hospital Erlangen, Germany

Abstract

Lymph nodes have high clinical relevance but detection is challenging as they are hard to see due to low contrast and irregular shape. In this paper, a method for fully automatic mediastinal lymph node detection in 3-D computed tomography (CT) images of the chest area is proposed. Discriminative learning is used to detect lymph nodes based on their appearance. Because lymph nodes can easily be confused with other structures, it is vital to incorporate as much anatomical knowledge as possible to achieve good detection rates. Here, a learned prior of the spatial distribution is proposed to model this knowledge. As atlas matching is generally inaccurate in the chest area because of anatomical variations, this prior is not learned in the space of a single atlas, but in the space of multiple ones that are attached to anatomical structures. During test, the priors are weighted and merged according to spatial distances. Cross-validation on 54 CT datasets showed that the prior based detector yields a true positive rate of 52.3% for seven false positives per volume image, which is about two times better than without a spatial prior.

1. Introduction

Lymph nodes play an important role in clinical practice, especially in the mediastinal area. They routinely need to be considered during oncological examination related to all kinds of cancer [6, 3], for instance lung cancer [10], where metastases settle in lymph nodes, but also lymphoma, which is a cancer of the lymphatic system itself. Furthermore, they are also relevant in case of inflammation in general.

Cancer causes affected lymph nodes to be enlarged. In order to assess the progress of the disease and to check whether treatment is effective, physicians are interested in statistics like the number of enlarged nodes or the total vol-
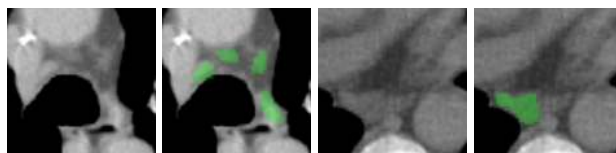

Figure 1. Two axial cross sections of CT volumes with expert-reviewed lymph node annotations (green).

ume of the nodes, but also in the spatial distribution, and changes over time. Patients are commonly examined using CT.

Manually counting and measuring lymph nodes in the images is not only cumbersome but also error prone because annotations from different human observers and even from the same human observer vary significantly. In practice, lymph nodes are not annotated individually because it would take too much time, though the clinical value would be high. An automatic detection is however challenging because lymph nodes have an attenuation coefficient similar to muscles and vessels and therefore low contrast to surrounding structures. Moreover, their shape and size varies a lot. Even a human needs days of training to consistently find lymph nodes in CT volume images. Examples of mediastinal lymph nodes are shown in Figure 1.

Up to now, there are few publications on automatic lymph node detection. In [9], two blob detectors which are called 3-D Min-DD filter and extended 3-D Min-DD filter are used in a cascade to detect lymph nodes in abdominal CT data. A Hessian based vessel detector, the CT Hounsfield units and morphological operations are used to reduce the number of false positives. In [7], a similar approach is used to detect lymph nodes in chest CT. Here, the first 3-D Min-DD filter is replaced with a Hessian based blob detector. The more expensive extended 3-D Min-DD filter is used at the second level of the cascade. For segmentation, a model based approach using mass spring models was proposed in [4]. It was also used for detection by plac-

1

ing models on a regular grid over the volume [5]. A lymph node was assumed at positions where the model fitting converged with a good score.

In this paper, we use discriminative learning techniques in combination with search space pruning based on [14] to detect lymph nodes from their appearance. The major contribution of this work is to combine this powerful detector with anatomical knowledge that is modeled as a spatial prior probability which is learned from annotated data. As inter-subject or atlas registration is generally inaccurate because of variations in the anatomy, we use a mixture of spatial priors that are attached to different anatomical structures, which are detected automatically.

The remainder of this paper is structured as follows: Section 2 explains the discriminative model used to recognize lymph nodes from their appearance. In section 3, the spatial prior model and its integration into the probabilistic detection framework is discussed. Section 4 presents results and section 5 concludes the paper.

2. Appearance based detection

As we are especially interested in the size and location, we want to find axis aligned bounding boxes of the lymph nodes. The boxes $\boldsymbol{b}$ are parameterized by their center $\boldsymbol{t}$ and size $\boldsymbol{s}$:

$$\boldsymbol{b} = (\boldsymbol{t}, \boldsymbol{s}) = (t_x, t_y, t_z, s_x, s_y, s_z) \qquad (1)$$

We use a multi-stage system of probabilistic boosting tree (PBT) classifiers [12]. These are binary decision trees with a strong Ada-Boost classifier at each node. In the first detection stage, a PBT classifier is trained with 3-D Haar-like features to learn the probability

$$p(m = 1 | \boldsymbol{T}(\boldsymbol{t})) \qquad (2)$$

of whether there is a lymph node model instance at a given position $\boldsymbol{t}$, which has similarities to the face detector proposed by Viola and Jones [13]. Here, $\boldsymbol{T}$ denotes the Haar-feature vector extracted at position $\boldsymbol{t}$. Haar features are used because they can be computed very efficiently so that it is even possible to search all positions in the volume exhaustively.

Given the output from the classifier, a set of position candidates $C_{T1} = \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{|C_{T1}|}\}$ is generated. If a fixed threshold θ is used and we select all $\boldsymbol{t}$ that satisfy $p(m = 1 | \boldsymbol{T}(\boldsymbol{t})) > \theta$, we run into the problem that lots of candidates are generated at lymph nodes which are clearly visible, but we don't get any candidates at lymph nodes which are hard to see. To overcome this, we use a technique proposed in [2]. First, a probability map is generated from the classifier output. This map is blurred using a Gaussian filter with a standard deviation of 1.5mm, and local maxima in the probability map are selected as candidates. An example of candidates extracted from the blurred probability
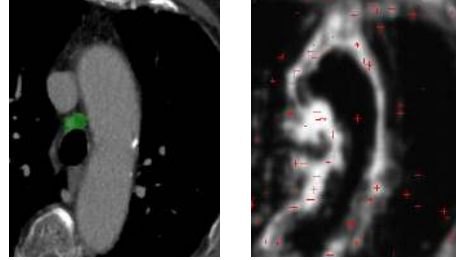

Figure 2. Local maxima of the probability map generated by the detector are used as position candidates. Note that this is a 2-D slice of a 3-D volume and points that look like local optima in 2-D are not necessarily local optima in 3-D.

map can be seen in Figure 2 along with the CT image data it was generated from.

Now, another PBT classifier is used to examine the position candidates in the set C_{T1} and to reject false positives, resulting in a set C_{T2} of candidates. The negative examples for the training phase of this classifier are generated by scanning images using the first classifier and collecting false positives.

So far, the size of the lymph nodes had not been taken into account. The classifiers for translation are trained with data of different sizes. We follow the idea of marginal space learning proposed in [14]: Instead of directly searching a high dimensional search space, which here consists of position $\boldsymbol{t}$ and size $\boldsymbol{s}$ parameters, candidates of position are generated using a detector that is trained on a margin of the search space, which is spanned by $\boldsymbol{t}$ in this case. A second detector trained to learn the probability $p(m = 1 | \boldsymbol{S}(\boldsymbol{t}, \boldsymbol{s}))$ of a lymph node at a given position with a given size then only needs to consider the position candidates C_{T2}, leading to an enormous speedup. Here, $\boldsymbol{S}(\boldsymbol{t}, \boldsymbol{s})$ denote so-called steerable features [14] evaluated at position $\boldsymbol{t}$ and size $\boldsymbol{s}$. These are simple features like the intensity, the gradient and nonlinear combinations evaluated on a regular grid of size $7 \times 7 \times 7$, scaled and translated according to $\boldsymbol{s}$ and $\boldsymbol{t}$.

The number of tree levels and the number of weak classifiers per Ada-Boost node are parameters of the PBT that allow to control overfitting. Here, we used trees with two levels and 20 weak classifiers per node.

3. Spatial prior

Up to now, only the image itself was used for detection and we did not exploit available prior knowledge. In particular, lymph nodes

- always lie in fat tissue, so space inside any organ can be excluded.

- In the remaining area, lymph nodes are not distributed equally. Instead, it is for instance much more likely to observe lymph nodes below the aortic arch and close to the trachea.

3.1. Automatic landmark detection and organ segmentation

In this work, this knowledge is represented as a spatial distribution of lymph nodes that is learned with respect to anatomical structures.

We automatically find a set of 20 salient anatomical landmarks, that are mostly but not exclusively in the chest area and can be detected robustly. The detection method used here is described in [11]. Examples of landmarks are the bifurcation of the trachea, the bottom tip of the shoulder blade left and right, the topmost point of the aortic arch and the topmost point of the lung left and right.

Besides the landmarks, a number of different organs are segmented. The lungs and the trachea are detected using simple thresholding followed by a morphological opening operation. The four heart chambers are segmented by fitting a model described in [14]. The esophagus is segmented using the approach of [8]. It is of special interest as it is often surrounded by lymph nodes, but at the same time can be confused with lymphatic tissue. All segmentation methods do not require user interaction.

3.2. Binary mask

We are looking for a spatial prior probability $p(m = 1|\boldsymbol{t})$ of observing a lymph node at a given location. A model for this prior is proposed that consists of three parts: The first part is a binary mask

$$B(\boldsymbol{t}) = \begin{cases} 0 & \text{if } \boldsymbol{t} \text{ is inside an organ} \\ 1 & \text{else} \end{cases} \quad (3)$$

that labels regions which cannot contain lymph nodes with 0 and other regions with 1. The lungs, the trachea, the esophagus and the heart are excluded, i. e. labeled with zero in the mask.

3.3. Global prior

The second part is a global soft prior

$$G(\boldsymbol{t}) \in [0, 1] \quad (4)$$

which is learned in the space of a reference patient. Non-rigid inter subject registration is used to map segmented lymph nodes from a set of test patients to the reference patient, where they are averaged. The segmentations are binary masks, and thus $G(\boldsymbol{t})$ is the spatial probability of lymphatic tissue. The learned prior is blurred with a Gaussian filter with a standard deviation of 12mm which is necessary because of limited training data.

The registration is based on the set of 20 landmarks. If a landmark is not detected, e. g. because it is not visible in the image, it is omitted. A thin-plate splines (TPS) transformation [1] is created from the detected landmarks and
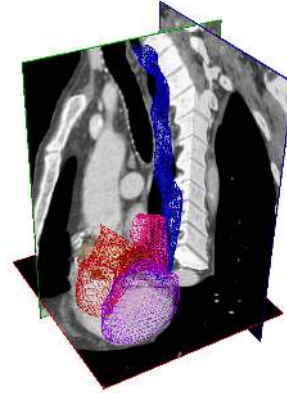

Figure 3. Heart and esophagus model fitted to a CT volume.

the reference landmarks and used for the warping. During training, the transformation maps from the reference space to the current image, and for the testing phase, it maps into the other direction.

3.4. Organ-specific local prior

The disadvantage of the global prior G is the limited accuracy which stems from inter-subject variations of anatomy and from missing landmarks on points or surfaces of interest. In general, the problem of finding a transformation f that maps from one image to another is ill-posed if certain points or regions, like for instance a gap between two organs, only exist in one, or if the relative position of two structures, e.g. the esophagus and the trachea, is different in the two images. This is a problem all atlas based techniques have to cope with.

In this work, a solution to this problem is proposed which is based on a mixture of spatially weighted priors. In addition to the landmark based global prior G, a set of local priors $L_i, i = 1 \ldots M$ is introduced. We already have the heart and the esophagus available, which are represented as triangular meshes. For the heart, the four chambers are treated like separate organs. The vertices of these meshes can be used as landmarks for a TPS based registration like in case of the global prior G. If they were simply added to the existing set of landmarks, however, the resulting TPS transformation would easily become rugged and produce folds or become inaccurate if it was constrained to be smooth because of the problems described above. Here, each segmented organ is associated with a prior L_i. It is learned in the space of the organ's mean shape, which is generated by averaging the points of a set of training shapes that were aligned by a generalized Procrustes analysis. Each local prior L_i is trained like the global, but now the vertices of the organ mesh serve as landmarks for the TPS transformation. In the testing phase, the final local prior $L(\boldsymbol{t})$ for a test volume is the sum of the local priors, weighted with $w_i(\boldsymbol{t})$

according to the minimum distance $d_i(\boldsymbol{t})$ of $\boldsymbol{t}$ to the surface of organ i:

$$L(\boldsymbol{t}) = \frac{1}{Z(\boldsymbol{t})} \sum_{i=1}^{M} L_i(\boldsymbol{t}) w_i(\boldsymbol{t}) \qquad (5)$$

$$Z(\boldsymbol{t}) = \sum_{i=1}^{M} w_i(\boldsymbol{t}) \qquad (6)$$

$$w_i(\boldsymbol{t}) = \begin{cases} \frac{(d_i(\boldsymbol{t}) - \theta)^2}{\theta^2} & \text{if } d_i(\boldsymbol{t}) < \theta \\ 0 & \text{else.} \end{cases} \qquad (7)$$

Here, θ denotes the maximal distance from the organ surface up to that the corresponding prior still has support. In the experiments, θ was set to 25mm. The weight w_i decreases with increasing distance because the TPS transformation is only accurate close to the surface.

3.5. Final prior and integration into detector

Now the binary mask B, the global prior G and the local prior L are merged into a final prior probability $p(m = 1|\boldsymbol{t})$ of observing a lymph node at position $\boldsymbol{t}$. It is modelled as

$$p(m = 1|\boldsymbol{t}) = B(\boldsymbol{t}) \left(w_{\max}(\boldsymbol{t}) L(\boldsymbol{t}) + (1 - w_{\max}(\boldsymbol{t})) G(\boldsymbol{t}) \right) \qquad (8)$$

where $w_{\max}$ is the maximum weight

$$w_{\max}(\boldsymbol{t}) = \max \left(w_1(\boldsymbol{t}), \ldots, w_M(\boldsymbol{t}) \right). \qquad (9)$$

During position detection, we are finally interested in the probability

$$p(m = 1|\boldsymbol{T}, \boldsymbol{t}) \qquad (10)$$

of whether there is a lymph node at a given position $\boldsymbol{t}$, with a given feature vector $\boldsymbol{T}$. With Bayes' rule, (10) can be reformulated as

$$p(m = 1|\boldsymbol{T}, \boldsymbol{t}) = \frac{p(\boldsymbol{T}, \boldsymbol{t}|m = 1)p(m = 1)}{p(\boldsymbol{T}, \boldsymbol{t})}. \qquad (11)$$

For simplification, we assume that the feature vector $\boldsymbol{T}$ is statistically independent from the position $\boldsymbol{t}$. This is an approximation as $\boldsymbol{T}$ obviously depends on $\boldsymbol{t}$, but $\boldsymbol{t}$ determines $\boldsymbol{T}$ only for a certain image. The assumption is justified by the fact that the spatial prior clearly improves the performance as we will see, which means that $\boldsymbol{T}$ does not contain much information about $\boldsymbol{t}$.

Now (11) may be transformed into

$$p(m = 1|\boldsymbol{T}, \boldsymbol{t}) = \frac{p(\boldsymbol{T}|m = 1)p(\boldsymbol{t}|m = 1)p(m = 1)}{p(\boldsymbol{T})p(\boldsymbol{t})} \qquad (12)$$

$$= \frac{p(m = 1|\boldsymbol{T})p(m = 1|\boldsymbol{t})}{p(m = 1)}, \qquad (13)$$
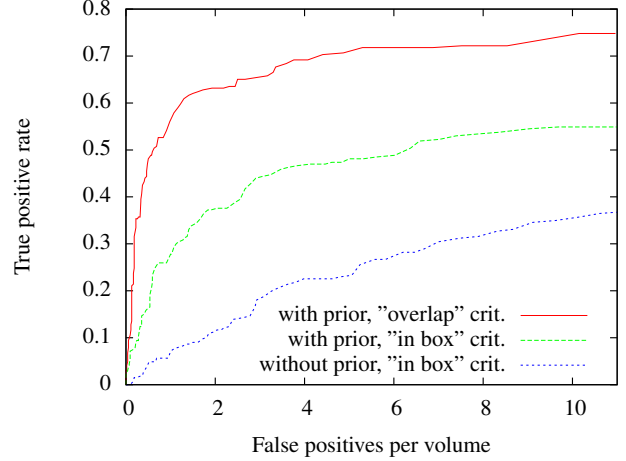

Figure 4. ROC curve comparing the performance of the proposed prior based method to the performance of only using the classifier output.

which is proportional to the product of (2) and (8) and is used as final translation detection score. A blurred map of this score is visualized in Figure 2 (right).

Similarly, if also scale is incorporated and we're interested in the probability of having a lymph node at a given position $\boldsymbol{t}$ with features $\boldsymbol{T}$ and $\boldsymbol{S}$, it can be expressed as

$$p(m = 1|\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{t}) = \frac{p(m = 1|\boldsymbol{S})p(m = 1|\boldsymbol{T})p(m = 1|\boldsymbol{t})}{p^2(m = 1)} \qquad (14)$$

under the assumption that the steerable features $\boldsymbol{S}$ are statistically independent from the Haar-like features $\boldsymbol{T}$ and position $\boldsymbol{t}$. This probability serves as final detection score. Based on it, a set C_{TS} of translation and scale candidates is generated. The best 400 candidates are spatially clustered and merged to generate the lymph node detections. Hierarchical clustering is used with a maximum distance of 12mm.

4. Results

In order to train the discriminative model and the spatial prior, 54 CT volume images of patients suffering from lymphoma were collected. The inter-slice spacing was 1mm, and the voxel spacing within an axial slice was in the range of 0.7mm to 0.9mm. All datasets were resampled to isotropic $1 \times 1 \times 1\text{mm}^3$ resolution. The mediastinal lymph nodes were manually segmented, and the segmentations were reviewed by an experienced radiologist. In total, 1086 lymph nodes were annotated.

The detection performance was evaluated using three-fold cross-validation. For each fold, both the spatial prior and the classifiers were trained on the training data and evaluated on the test data. The classifiers were only trained on lymph nodes that have a minimum size of 10mm in at least two dimensions. Smaller lymph nodes are usually not

Method	Body region	TP criterion	num. vol.	size/mm	TP	FP	FN	TPR	FP per vol.
Kitasaka *et al.* [9]	Abdomen	overlap	5	> 5.0	126	290	95	57.0%	58
Feuerstein *et al.* [7]	Mediastinum	overlap	5	> 1.5	87	567	19	82.1%	113
This method	Mediastinum	overlap	54	>10.0	174	157	92	65.4%	2.9
Dornheim [5]	Neck	unknown	1	> 8.0	29	9	0	100%	9
This method	Mediastinum	in box	54	>10.0	117	157	149	44.0%	2.9
This method	Mediastinum	in box	54	>10.0	139	379	127	52.3%	7.0
Intra-observer var.	Mediastinum	in box	10	>10.0	23	8	19	54.8%	0.8

Table 1. Detection results compared to state of the art methods and to the intra-observer variability. In the top and bottom three rows, a ground truth lymph node is considered as detected if it overlaps with a detection, and if the center of a detection falls within the tight bounding box of the lymph node, respectively.
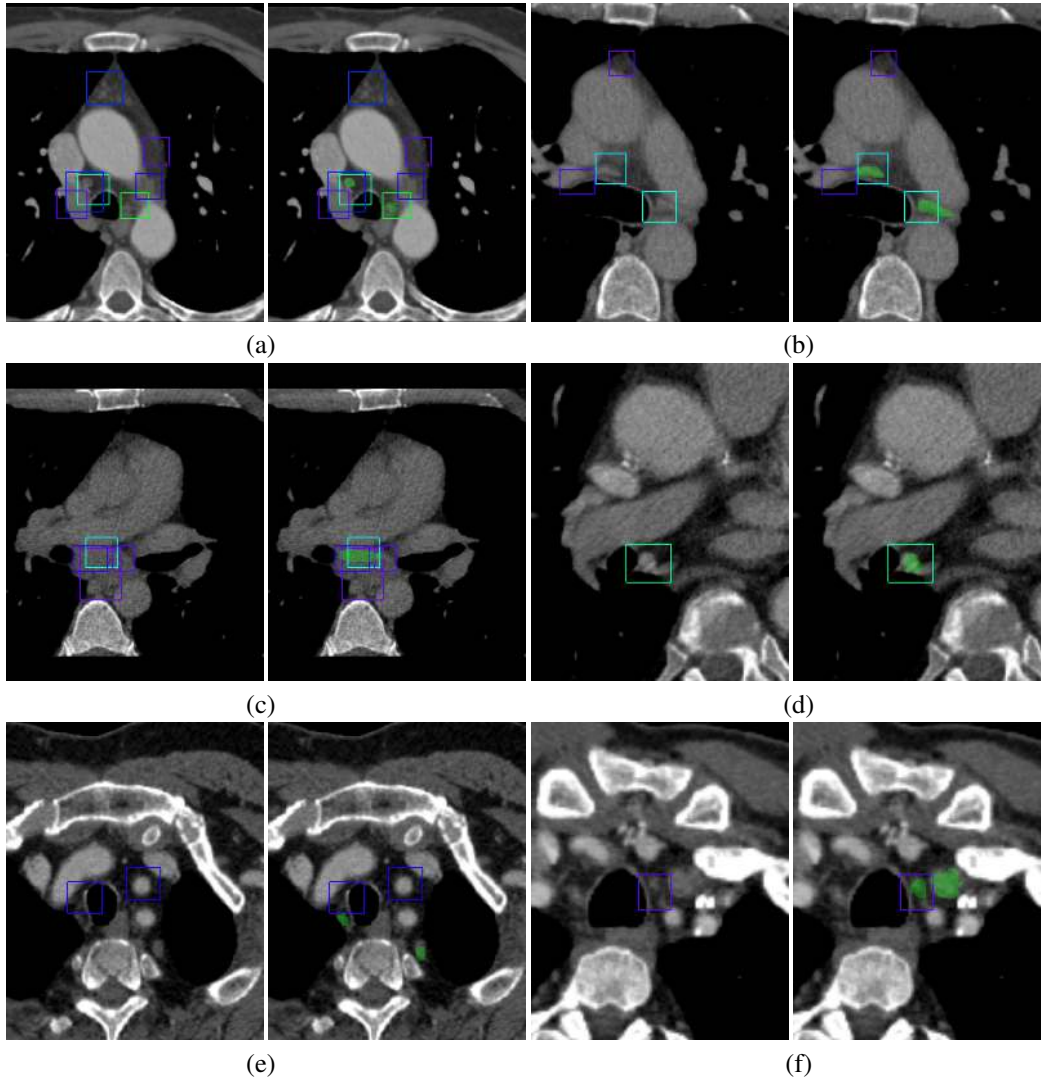

Figure 7. Examples of detection results on unseen data, shown in 2-D, with (even columns) and without (odd columns) ground truth segmentation in green. The confidence $p(m = 1|S, T, t)$ of a detection is color coded. Violet means lowest, turquoise is medium and green is high confidence. (a-d): The lymph node were detected. There are false positives, but the true positives have higher confidence. (e-f): Examples of false negatives.

pathologic [3] and were therefore neglected. The set of manual segmentations contained six huge cases with a size exceeding 5cm which were mostly not single nodes but a cluster of lymph nodes which were densely packed so that
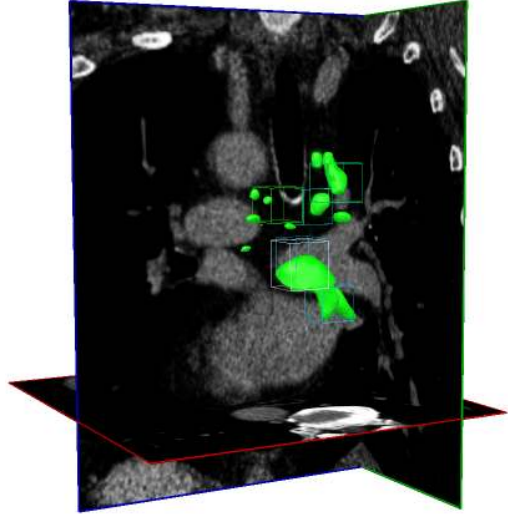
Figure 5. Detection results on unseen data shown in 3-D along with ground truth segmentations in green.
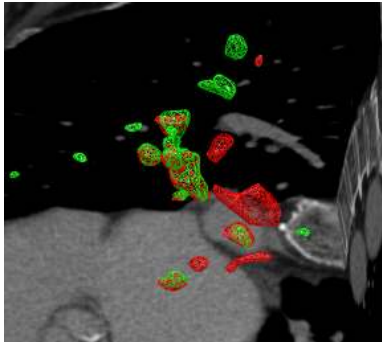

Figure 6. Two manual segmentations from the same person.

axial   sagittal   coronal
Figure 8. Final prior computed for a test volume.

the boundaries become invisible. Such cases were removed from the training set in order not to distract the detector with few extreme examples. Among the segmented lymph nodes, 266 were used for training. In order to achieve a better generalization and to avoid overfitting, the training data was mirrored by all three coordinate planes, resulting in $2^3 = 8$ times more training examples. For testing, only the original data was used.

During test, two different criteria were used to separate true positives (TP) from false positives (FP). In the first case, a manually segmented lymph node is considered as detected if there is overlap between the segmentation and the detection. This criterion was used in [9] and [7] according to the authors. Here, it is called "overlap". This error measure is however a suboptimal choice because a single huge detection covering the whole volume would result in a true positive rate (TPR) of 100% with zero FP, although the detection is obviously meaningless. Therefore, we measured the performance also with a second criterion called "in box", which considers a lymph node as detected if the center of a detected box is inside the tight axis-aligned bounding box of the lymph node. A lymph node was considered as false negative (FN) if its size was at least 10mm and it was not detected.

The performance of the detectors considerably depends on how the negative training examples are selected. Instead of randomly sampling images, the position samples for the first detector in the chain were drawn only from the regions of the binary mask B labeled with one. This avoids confusing the detector with data it won't ever see in the testing phase.

A ROC curve of the detection performance is shown in Figure 4. For example, we get a TPR of 22.6% with 4.0 FP per volume on average for the "in box" error measure if we only rely on the score of the classifiers. If also the spatial prior is used, we get a TPR of 46.6% with 3.8 TP per volume, which is more than two times better. The figure also shows the results of the prior based method for the "overlap" error measure.

In Table 1, the detection performance is compared to state of the art methods and to the performance of a human observer. In [9], a TPR of 57% was reported for abdominal lymph nodes, but with 58 false alarms per volume. In [7], 82.1% of mediastinal lymph node were detected with 113 FP per volume. When the same "overlap" error measure was used, the method proposed here detected 65.4% of the lymph nodes with only 2.9 FP per volume. In [5], a TPR of 100% is reported with 9 FP, but it was evaluated on a single volume image only, and regions of interest were selected manually. The comparability with [9, 7, 5] is however limited because of different minimum sizes above which missed lymph nodes are counted as false negatives, different datasets and annotations, the "overlap" error measure and partially different body regions.

In order to compare the automatic detection results with the performance of a human, we did an experiment on the intra-human observer variability. Ten of the CT volumes were annotated a second time by the same person a few months later. The first segmentations served as ground truth, and the second ones were considered as detections. TPR and FP were measured in the same way as for the automatic detection. The TPR was 54.8% with 0.8 false positives per volume on average. While 0.8 FP is very low, a TPR of 54.8% shows that finding lymph nodes in CT is quite challenging also for humans. Figure 6 shows the first

Landmark detection	heart and esophagus segmentation	prior computation	detection	**total**
5.86	20.7	66.3	42.0	**134.8**

Table 2. Computation time in seconds for the different steps of the proposed method.

and the second segmentations for one of the ten datasets.

In Figure 8, orthogonal cross-sections of the final spatial prior (8) for a test volume are shown. The location of lymph nodes is highly constrained by the prior. It was learned that lymph nodes are more likely close to the trachea, especially close to its bifurcation, and also close to the esophagus, which is in fact true as these are natural gateways to the body.

Figure 7 shows some detection examples. In general, the detection score, which is color coded, reflects well the quality of the detection. False positives are mostly structures that look similar to lymph nodes like vessels, or sometimes they are at locations where lymph nodes are common. One reason for this behavior is the spatial prior, and another reason is that the detectors also learned common surrounding structures of lymph nodes. The technique of generating more training data by mirroring helps to lower this effect. Figure 5 shows detection results in 3-D.

The computational requirements were measured for a volume of size $183 \times 179 \times 251$ voxel with a voxel spacing of $1 \times 1 \times 1$ mm^3. The results are shown in Table 2. Including landmark detection and segmentation of the heart and the esophagus, lymph node detection takes 2min 15s on a 2.2GHz dual core processor. The method of [7] takes 1-6min, in [5], a computation time of 17min is reported, and the method of [9] takes 2-3h.

5. Discussion

This work presents a method for fully automatic detection of mediastinal lymph nodes, which is a challenging problem due to low contrast to surrounding structures and clutter. In contrast to prior work, we heavily rely on prior anatomical knowledge, which is modelled as a spatial prior probability and learned from annotated data. To handle anatomical variations, a mixture of local and global priors is proposed, which are non-rigidly attached to segmented organs and landmarks. This is combined with a powerful detector. Evaluation on 54 CT volume images showed competitive results and the advantage of a prior based approach. The method can easily be adapted to other regions of the body, and also to other node-like structures and lesions, for example tumor metastases.

References

[1] F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *PAMI*, 11(6):567–585, 1989. 3

[2] T. Chen, W. Zhang, S. Good, K. S. Zhou, and D. Comaniciu. Automatic ovarian follicle quantification from 3d ultrasound data using global/local context with database guided segmentation. In *ICCV*, 2009. 2

[3] A. J. de Langen, P. Raijmakers, I. Riphagen, M. A. Paul, and O. S. Hoekstra. The size of mediastinal lymph nodes and its relation with metastatic involvement: a meta-analysis. *Eur J Cardiothorac Surg*, 29(1):26–29, 2006. 1, 5

[4] J. Dornheim, H. Seim, B. Preim, I. Hertel, and G. Strauß. Segmentation of neck lymph nodes in ct datasets with stable 3d mass-spring models. In *MICCAI (2)*, pages 904–911, 2006. 1

[5] L. Dornheim and J. Dornheim. Automatische detektion von lymphknoten in ct-datensätzen des halses. In *Bildverarbeitung für die Medizin*, pages 308–312, 2008. 2, 5, 6, 7

[6] B. V. Duwe, D. H. Sterman, and A. I. Musani. Tumors of the mediastinum. *Chest*, 128(4):2893–2909, 2005. 1

[7] M. Feuerstein, D. Deguchi, T. Kitasaka, S. Iwano, K. Imaizumi, Y. Hasegawa, Y. Suenaga, and K. Mori. Automatic mediastinal lymph node detection in chest ct. In *SPIE Medical Imaging*, Orlando, Florida, USA, February 2009. 1, 5, 6, 7

[8] J. Feulner, S. K. Zhou, A. Cavallaro, S. Seifert, J. Hornegger, and D. Comaniciu. Fast Automatic Segmentation of the Esophagus from 3D CT Data Using a Probabilistic Model. In *MICCAI(1), LNCS 5761*, pages 255–262, 2009. 3

[9] T. Kitasaka, Y. Tsujimura, Y. Nakamura, K. Mori, Y. Suenaga, M. Ito, and S. Nawano. Automated extraction of lymph nodes from 3-d abdominal ct images using 3-d minimum directional difference filter. In *MICCAI (2)*, pages 336–343, 2007. 1, 5, 6, 7

[10] T. C. McLoud, P. M. Bourgouin, R. W. Greenberg, J. P. Kosiuk, P. A. Templeton, J. A. Shepard, E. H. Moore, J. C. Wain, D. J. Mathisen, and H. C. Grillo. Bronchogenic carcinoma: analysis of staging in the mediastinum with CT by correlative lymph node mapping and sampling. *Radiology*, 182(2):319–323, 1992. 1

[11] S. Seifert, A. Barbu, S. K. Zhou, D. Liu, J. Feulner, M. Huber, M. Suehling, A. Cavallaro, and D. Comaniciu. Hierarchical parsing and semantic navigation of full body ct data. In J. P. W. Pluim and B. M. Dawant, editors, *Medical Imaging 2009: Image Processing*, volume 7259, page 725902. SPIE, 2009. 3

[12] Z. Tu. Probabilistic boosting-tree: learning discriminative models for classification, recognition, and clustering. *ICCV*, 2:1589–1596, 2005. 2

[13] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *CVPR*, 1:511, 2001. 2

[14] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Fast automatic heart chamber segmentation from 3d ct data using marginal space learning and steerable features. *ICCV*, pages 1–8, Oct. 2007. 2, 3