Decision Forests with Long-Range Spatial Context for Organ Localization in CT Volumes

A. Criminisi, J. Shotton, and S. Bucciarelli

Microsoft Research Ltd, Cambridge, UK

Abstract. This paper introduces a new, efficient, probabilistic algorithm for the automatic analysis of 3D medical images. Given an input CT volume our algorithm automatically detects and localizes the anatomical structures within, accurately and efficiently.
Our technique builds upon randomized decision forests, which are enjoying much success in the machine learning and computer vision communities. Decision forests are enriched here with learned visual features which capture long-range spatial context. In this paper we focus on the detection of human organs, but our general-purpose classifier might be trained instead to detect anomalies. Applications include (and are not limited to) efficient visualization and navigation through 3D medical scans.
The output of our algorithm is probabilistic thus enabling the modeling of uncertainty as well as fusion of multiple sources of information (e.g. multiple modalities). The high level of generalization offered by decision forests yields accurate posterior probabilities for the localization of the structures of interest. High computational efficiency is achieved thanks both to the massive level of parallelism of the classifier as well as the use of integral volumes for feature extraction.
The validity of our method is assessed quantitatively on a ground-truth database which has been sanitized by medical experts.

1 Introduction

This paper presents a new, efficient algorithm for the accurate detection and localization of anatomical structures within CT scans. This work represents a significant step towards automatic parsing and understanding of medical images.

Our effort is motivated by recent studies which indicate how the great majority of a radiologist's time is spent searching through scanned volumes (often slice by slice) and navigating through visual data. Even with modern 3D visualization tools locating the organ(s) of interest and selecting optimal views is time consuming. Automatic tools for localizing major anatomical structures within 3D scans promises to speed up navigation and improve the user's work-flow [1]. For instance, a cardiologist may just click on a button to take him/her to the most appropriate view of the heart and its valves. Robust and efficient, probabilistic organ detection is also useful as input to other, more specialized tasks, e.g. detecting the heart to initialize a coronary tree tracer tool.

The two main contributions are: 1) We introduce an efficient algorithm for organ detection and localization which negates the need for atlas registration;

thus overcoming issues related to, e.g. : i) possible lack of atlases, and ii) selecting the optimal model for geometric registration. 2) We introduce new, context-rich visual features which capture long-range spatial correlations efficiently. The simplicity of our features combined with the intrinsic parallelism of our classifier yield high computational efficiency. Finally, our algorithm produces probabilistic output, useful for instance to keep track of uncertainty in the results, to take into account prior information (e.g. about global location of organs) or to fuse multiple sources of information (e.g. different acquisition modalities).

The proposed algorithm is applied here to the task of localizing nine anatomical structures (`head`, `heart`, `left eye`, `right eye`, `l. kidney`, `r. kidney`, `l. lung`, `r. lung`, and `liver`) in CT volumes with varying resolution, varying cropping, different patients, different scanner types and settings, contrast enhanced and not etc. Quantitative assessment is executed on a number of manually labelled ground-truth CT volumes.

Previous work. In the last few years research in object detection and recognition has made huge progress. The published work which is relevant to medical applications may be broadly categorized into the following three groups:

Geometric methods include template matching, and convolution techniques [2]. Geometrically meaningful features are used in [3, 4] for the segmentation of the aorta and the airway tree, respectively. Such geometric approaches often have problems capturing invariance with respect to deformations (e.g. due to pathologies), changes in viewing geometry (e.g. cropping) and changes in intensity. Techniques built upon "softer" geometric models with *learned* spatial correlations have been demonstrated to work well both for rigid and deformable objects [5].

Atlas-based techniques have enjoyed much popularity. Recent techniques for single and multiple organ detection and segmentation based on the use of probabilistic atlases include [6–10]. The apparent conceptual simplicity of such algorithms is in contrast to the need for accurate, deformable registration algorithms. The major problem with n-dimensional registration is in selecting the appropriate number of degrees of freedom of the underlying geometric transformation; especially as it depends on the level of rigidity of each organ/tissue.

Supervised, discriminative classification. Discriminative classification algorithms such as Support Vector Machines (SVM), AdaBoost and Probabilistic Boosting Trees have been applied successfully to tasks such as: automatic detection of tumors [11–14], pulmonary emphysema [15], organs in whole-body scans [19] and brain segmentation [16–18]. Our approach is also a discriminative classification technique. It achieves multi-class recognition efficiently and probabilistically. The classifier employed here is a random decision forest which, in non-medical domains has been shown to be better suited to multi-class problems than SVMs, as well as being more effective than boosting [20, 21]. A model of spatial context is learned by automatically selecting visual features which capture the relative position of visual patterns. Next we describe the details of our technique.

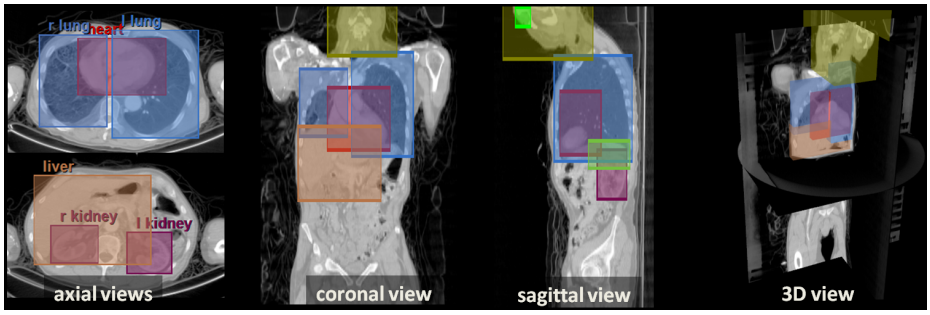

Fig. 1. Constructing labelled ground-truth databases. Organs within 3D CT scans are labelled via 3D, axis-aligned bounding boxes; different colours indicating different organs. Note that the fact that the boxes overlap is not a problem as they are used to indicate the position of the organ centre and the organ's approximate extent.

2 Automatic Parsing of CT Volumes

This section presents our ground-truth database, describes the decision forest classifier in the context of CT images and illustrates the visual features employed.

2.1 Labelled ground-truth database and exemplars

We have 39 CT volumes which have been annotated with 3D bounding boxes centred on each organ using our own annotator tool (shown in fig. 1). The user loads a CT scan, locates the organ of interest and draws a 3D box tightly around the organ. The database is split randomly into training and test sets as outlined in section 3. We focus on the following nine organs: `head`, `heart`, `left eye`, `right eye`, `l. kidney`, `r. kidney`, `l. lung`, `r. lung`, and `liver`. The use of axis-aligned boxes enables speedy manual annotation and is sufficient for tasks such as detection[1]. Our dataset comprises both contrasted and non-contrasted CT data, from different patients, cropped in different ways, with different resolutions and acquired from different scanners.

The goal is to determine the centre of each organ in previously unseen CT scans. A supervised technique such as ours needs to be trained from positive and negative examples of organ centres. Exemplars are provided from the annotation boxes as follows (cf. fig. 2). For each organ (e.g. the right kidney in fig. 2) we denote its annotation box with B^a. The set of positive training example points for the organ centre are defined as the set of points *within* a small box B^+; with B^+ of fixed size and located in the centre of B^a. Similarly, negative examples are all points *outside* the box B^- with same centre and aspect ratio as B^a but 50% in size. The region between B^- and B^+ is ignored.

[1] 2D annotation boxes (with no pixel-wise annotation) are used extensively in the PASCAL VOC challenges: `pascallin.ecs.soton.ac.uk/challenges/VOC/voc2009/`

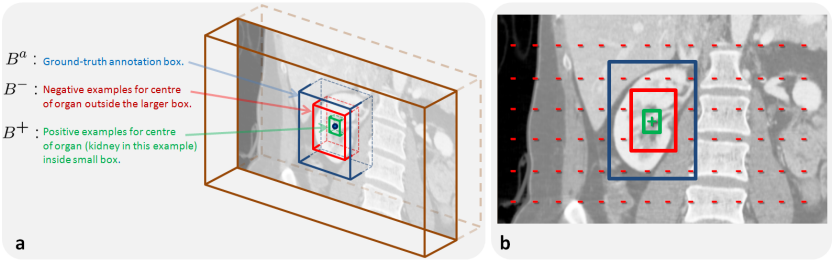

Fig. 2. Positive and negative training examples. (a) A 3D view illustrating the ground truth annotation box B^a, the box of positive examples B^+ and the box of negative examples B^-. **(a)** A 2D view further clarifying the regions where exemplars for the organ centre are taken. Positive examples are sampled within the box B^+ (in green). Negative example points are sampled outside the box B^- (in red).

2.2 Decision forests for recognition in 3D medical images

This section describes our adaptation of decision forests to the task of organ detection and localization in 3D CT scans.

A random decision forest [23, 24] is a collection of deterministic decision trees. Decision *trees* are popular classification algorithms which are known to suffer from over-fitting (poor generalization). Recently, it has been shown that the ensemble of many randomly trained decision trees (a random forest) yields much better generalization while maintaining the advantages of conventional decision trees [23]. Intuitively, where one tree fails the others do well.

We use the following notation. A voxel in a volume V is defined by its coordinates $\mathbf{x} = (x, y, z)$. The forest is composed of T trees denoted $\Psi_1, \cdots, \Psi_t, \cdots, \Psi_T$; with t indexing each tree (fig. 3). In each tree, each internal node (split node) performs a binary test on the input data and based on the result directs the data to the left or right child. The leaf nodes do not perform any action, they just store probability distributions over the organs of interest. Next we describe how the split functions are chosen and how the leaf probabilities are computed.

Forest training. Each point $\mathbf{x}$ of each *training* volume is associated with a known (manually obtained) class label $Y(\mathbf{x})$. The label indicates whether the point $\mathbf{x}$ belongs to the positive set of organ *centres* (see fig. 2) or not. Thus, $Y(\mathbf{x}) \in \{$ `head, heart, left eye, right eye, l. kidney, r. kidney, l. lung, r. lung, liver, background` $\}$, where the `background` label indicates that the point $\mathbf{x}$ is not an organ centre.

During training T is fixed (we use $T = 10$). Then, each point $\mathbf{x}$ is pushed through each of the trees starting at the root. Each split node applies the following binary test: $\xi > f(\mathbf{x}; \boldsymbol{\theta}) > \tau$ and sends the data to the respective child node. $f(\cdot)$ is a function applied to $\mathbf{x}$ with parameters $\boldsymbol{\theta}$. The parameters $\boldsymbol{\theta}$ identify the visual features which needs be computed. Features are described in the next section; for now it suffices to say that f computes some scalar filter response at $\mathbf{x}$. ξ and τ are parameters of the split node. The purpose of training is to optimize

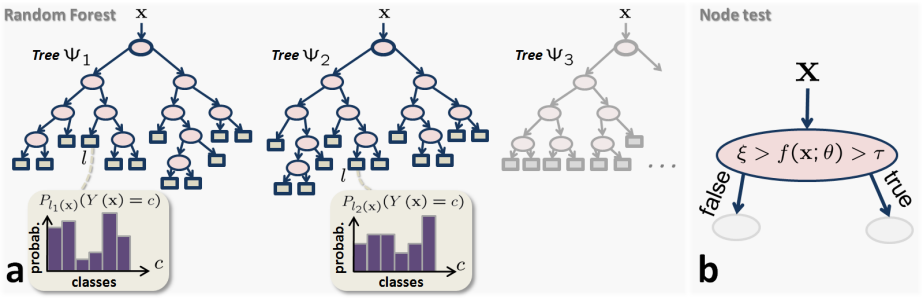

Fig. 3. Random Decision Forests. (a) An example random decision forest made of 3 trees. In each tree the internal nodes (shown with ellipses) perform simple tests on the input data while the leaf nodes (shown as squares) store the posterior probabilities over the classes being trained. During testing, a data point is pushed simultaneously through all T trees until it reaches T leaf nodes. The probability assigned to that point is the average of the probabilities of all the reached leaves (see text). **(b)** Each internal node performs a simple binary tests on the input data $\mathbf{x}$, based on the feature response $f(\mathbf{x};\boldsymbol{\theta})$. The quantities ξ, τ and $\boldsymbol{\theta}$ are parameters of the splitting test in that node.

the values of $\boldsymbol{\theta}, \xi, \tau$ of each split node by maximizing the data information gain, just like in the standard C4.5 tree training algorithm [25].

Injecting randomness for improved generalization. However, unlike standard tree training methods, here the parameters of each split node are optimized only over a randomly sampled subset Θ of all possible features (here $|\Theta| = 500$, details in section 2.3). This is an effective and simple way of injecting randomness into the trees, and it has been shown to improve generalization.

During node optimization all available features $\boldsymbol{\theta}_i \in \Theta$ are tried one after the other, in combination with many discrete values for the parameters ξ and τ. The combination $\xi^*, \tau^*, \boldsymbol{\theta}^*$ corresponding to the maximum information gain is then stored in the node for future use. The expansion of a node is stopped when the maximum information gain is below a fixed threshold. This gives rise to asymmetrical trees which naturally stop growing when no further nodes are needed. In this work the maximum tree depth D is fixed at $D = 15$ levels.

Finally, by simply counting the labels of all the training points which reach each leaf node we can associate each tree leaf with the empirical distributions over classes $P_{l_t(\mathbf{x})}(Y(\mathbf{x}) = c)$, where l_t indexes the leaf node in the t^{th} tree (fig. 4f). This training procedure is repeated for all T component trees.

Testing. During testing each point $\mathbf{x}$ of a previously unseen CT volume is simultaneously pushed through each of the T trees until it reaches a leaf node. Thus, the same input point $\mathbf{x}$ will end up in T different leaf nodes, each associated with a different posterior probability. The output of the forest, for the point $\mathbf{x}$, is the mean of all such posteriors, i.e. :

$$P(Y(\mathbf{x}) = c) \quad = \quad \frac{1}{T}\sum_{t=1}^{T} P_{l_t(\mathbf{x})}(Y(\mathbf{x}) = c)\,. \tag{1}$$

Other ways of combining the tree posteriors have been explored and simple averaging appears to be the most effective (as demonstrated also in the vast literature). Also, analyzing the variability of individual tree posteriors carries useful information about the uncertainty of the final forest posterior.
Organ detection. At this point detecting the presence/absence of an organ c is done simply by looking at the max probability $P_c = \max_{\mathbf{x}} P\left(Y(\mathbf{x}) = c\right)$. The organ c is considered present in the volume if $P_c > \beta$, with $\beta = 0.5$.
Organ localization. The centre of the organ c is estimated by marginalization over the volume V:

$$\mathbf{x}_c = \int_V \mathbf{x} \;\; p(\mathbf{x}|c) \;\; d\mathbf{x}, \tag{2}$$

where the likelihood $p(\mathbf{x}|c) = P\left(Y(\mathbf{x}) = c\right)$ by using Bayes rule and assuming uniform[2] distribution for organs. Furthermore, maximum a-posteriori classification for each voxel $\mathbf{x}$ may also be obtained as: $c^* = \arg\max_c P\left(Y(\mathbf{x}) = c\right)$. After having described our classification algorithm, next we provide details of the visual features employed.

2.3 Visual features and learned spatial context

The problem with identifying anatomical structures in CT images is that different organs may share similar intensity values. Thus, local intensity information is not sufficiently discriminative and further information such as texture, spatial context and topological cues must be used to have any chance of success. The problem then is how to capture and model such information efficiently.

Here we consider visual features which capture both the appearance of anatomical structures as well as their relative position (context) within the decision forest framework. For each location $\mathbf{x}$ context is modeled by integrating information coming from multiple regions which are offset by a quantity $\boldsymbol{\Delta}$ in a given direction. Figures 4 explains the main concepts with a 2D illustration. A feature $\boldsymbol{\theta}$ is defined as a reference point $\mathbf{o}$ paired with two boxes F_1, F_2 and two signal channels C_1, C_2. The shapes F_i are just 3D boxes displaced with respect to $\mathbf{o}$. The channels C_i could be for example the CT intensity ($C(\mathbf{x}) = I(\mathbf{x})$), or the magnitude of the 3D gradient ($C(\mathbf{x}) = |\nabla I(\mathbf{x})|$). Given a point $\mathbf{x}$ in a volume, computing the feature response $f(\mathbf{x}; \boldsymbol{\theta})$ corresponds to aligning the reference point $\mathbf{o}$ of the feature $\boldsymbol{\theta}$ with the point $\mathbf{x}$ and computing $f(\mathbf{x}; \boldsymbol{\theta}) = \sum_{\mathbf{q} \in F_1} C_1(\mathbf{q}) - b \sum_{\mathbf{q} \in F_2} C_2(\mathbf{q})$. The parameter $b \in \{0, 1\}$ indicates whether both feature boxes are used or only one (in fig. 4 $b = 0$ for simplicity).

As shown in fig. 4 these features tends to capture the relative layout of visual patterns (e.g. kidney patterns tend to occur a certain distance away, in a certain direction, from liver patterns, fig. 4d). The use of rectangular regions enables efficient *integral volume* processing [29, 30, 16]. Our features may be thought of as a generalization of the Haar-like features used in [26, 30, 16, 17]. In fact, we do not use manually predefined Haar subdivisions of a canonical cuboid. Our classifier is free to select features with very large offsets $\boldsymbol{\Delta}$, which enables

[2] Alternatively one can weight each class based on its own volume in the training set

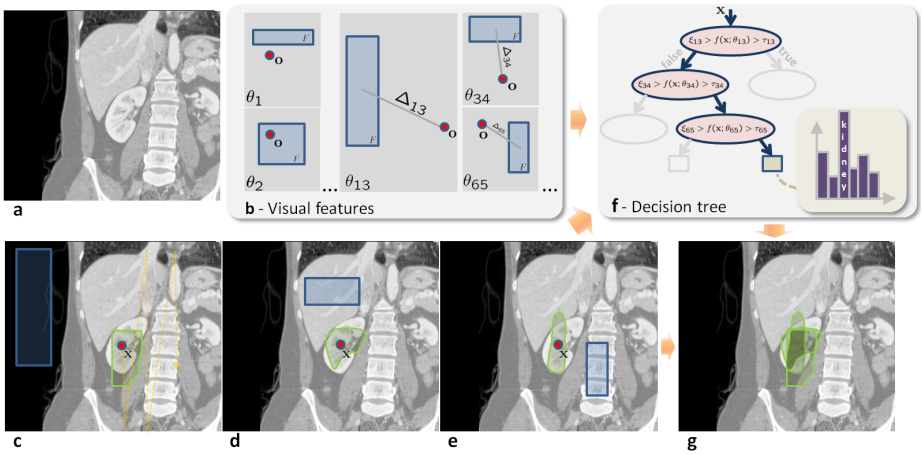

Fig. 4. Context-rich visual features, a 2D illustration. (a) Coronal view of a patient's abdomen. **(b)** Features (denoted $\boldsymbol{\theta}_i$) are defined as the rigid pairing between a box F and a reference point $\mathbf{o}$. Here we show only some of the infinite possible features. In practice we use 3D axis-aligned boxes. **(c)** Computing the feature response $f(\mathbf{x};\boldsymbol{\theta})$ at position $\mathbf{x}$ within a volume corresponds to aligning $\mathbf{o}$ with $\mathbf{x}$ and computing the sum $f(\mathbf{x};\boldsymbol{\theta}) = \sum_{\mathbf{q}\in F} I(\mathbf{q})$ (*cf.* text. For simplicity here we use intensity as the channel and only one rectangle). For feature $\boldsymbol{\theta}_{13}$ when $\mathbf{o}$ is on the kidney the rectangle F is in a region of low density (air). Thus the value of $f(\mathbf{x};\boldsymbol{\theta})$ is small for those points. During training the algorithm will learn that feature $\boldsymbol{\theta}_{13}$ is discriminative for the position of the right kidney when associated with a small, positive value of the threshold ξ_{13} (with $\tau_{13} = -\infty$). The region for which the condition $\xi_{13} > f(\mathbf{x};\boldsymbol{\theta}) > \tau_{13}$ is true is shown in green. **(d,e)** Similar to (c) but with different features. **(f)** Training associates each node with optimal values of $\xi, \tau, \boldsymbol{\theta}$. In this example, a data point which follows the highlighted path (in blue) gets assigned a high probability of being the centre of a kidney. **(g)** The points which satisfy all three conditions in (f) lie in the intersection of the three regions (c, d, e), highlighted in dark green, inside the organ of interest.

capturing very *long-range* spatial interactions. Inspection of the trained trees reveals that often the Δ of selected features can be as large as the image width. For simplicity, in this paper we only consider intensity and gradient as channels. However, our features are more flexible and general than that as they allow to incorporate complex filters such as SIFT, HOG etc. Multiple modalities may also be exploited; e.g. in the case of MR one may use T1, T2, FLAIR etc. More complex visual cues such as the ones described in [27, 28] or differently shaped aggregation regions may also be employed.

During training, for each split node the set Θ is obtained by randomly generating for each feature the two boxes F_1, F_2 (e.g. their centre and dimensions are randomly selected) and the corresponding channels C_1, C_2. Then all nodes are optimized and once training completes the trees, their nodes and the selected features are frozen and the testing phase proceeds deterministically.

2.4 Discussion and comparisons

The classifier used here is related to the Probabilistic Boosting Tree in [16]. In our case, the tree nodes contain test functions that are simpler than the boosters used in [16], with advantages in terms of speed both during training and testing. Furthermore, as shown in [20], a collection of simple, randomized trees tends to yield better generalization than a single tree of boosters.

In [17] the authors capture context by means of an algorithm which at each iteration uses the posteriors of the previous iteration as features. This produces good results at a cost of multiple iterations. Our algorithm is not iterative and captures spatial correlations of visual patterns, namely "appearance context". Furthermore, our kernels have much longer range. Finally, we do not require preregistration of the CT volumes.

Localizing anatomical structures by atlas registration is a popular option. However, such techniques have to deal with issues such as: i) the optimal choice of degrees of freedom of the registration model (e.g. both fully rigid and fully deformable transformations are bad); ii) the optimal choice of the reference template (e.g. an adult male body? a child? or a woman? contrast enhanced or not?); and iii) robustness to anatomical anomalies (training a classifier on data which presents anomalies allows the system to learn invariance to those).

The work in [19] makes use of information gain to optimize the scheduling of single-organ boosted detectors. In our work we use information gain at the level of feature selection, and detection happens via an ensemble of decision trees simultaneously for all organs. The selected features are organized hierarchically, with the most discriminative ones in the top layers of each tree. This has the advantage of "sharing" the most discriminative features amongst classes (organs) and sets of classes, with positive effects on generalization (e.g. see [31] for details on feature sharing and [20] for a detailed comparison between AdaBoost, decision trees and decision forests). Next we quantify the performance of the proposed algorithm and compare it to some known alternatives.

3 Experimental Results and Validation

This section presents qualitative and quantitative assessment of the accuracy of our algorithm applied to the tasks of organ detection and localization.

3.1 Automatic organ detection and localization

Qualitative results are shown in fig. 5. Our classifier applied to previously unseen CT scans produces accurate posteriors for the location of organ centres. In these visualizations the computed posteriors are used to modulate the transfer function employed during 3D rendering. For instance, notice how the mass of the heart probability (in red) is correctly concentrated around the centre of the heart region. Similarly for the light brown region indicating the liver, etc.
Quantitative evaluation of accuracy. Localization accuracy is assessed here by running training and testing multiple times. In each round the database is split

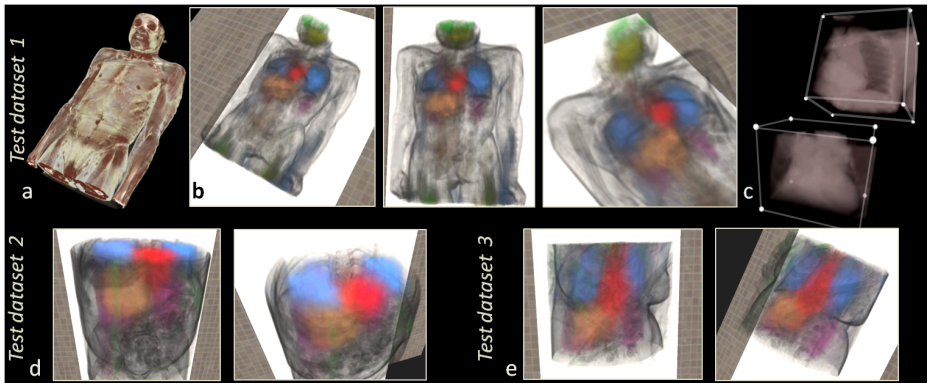

Fig. 5. Results of automatic organ detection and localization. (**a**) The original 3D CT data rendered using a manually-designed colour transfer function. (**b**) Three views of the 3D organ posterior probabilities computed by our algorithm for the localization problem. Different colours indicate different organs. Larger opacities indicate larger probability of a voxel being the organ centre. Notice how well eyes (green), head (yellow), heart, lungs, liver and even kidneys (purple) have been localized. A faint body outline is shown to aid visualization. (**c**) 3D views of the automatically detected bounding boxes including the heart and left lung. (**d,e**) Results on two more test datasets. The different datasets (related to different patients) are cropped differently and have different resolutions.

randomly into a training and a test set (with approximate ratio of 2 : 1). For all algorithms evaluated in this section the same 10 random splits are used. The forest is optimized on the training set only, and then applied on the test set. Then, the location of each organ centre is computed and compared with ground truth. Resulting localization errors collected from 10 runs (with $D = 15, |\Theta| = 500$) are shown below (in mm).

organ	`head`	`heart`	`l. eye`	`r. eye`	`l. kidney`	`r. kidney`	`l. lung`	`r. lung`	`liver`	*mean across organs*
median	25.58	18.31	24.04	25.71	13.52	29.49	22.93	21.94	19.01	**22.28 mm**
mean	29.92	21.32	28.78	27.14	25.42	44.52	27.05	26.75	22.68	28.18 mm
std	12.80	5.67	18.88	18.66	9.82	15.00	7.25	9.44	5.30	11.42 mm

Standard deviations (computed across the means of all runs) are reported here only to provide an indication of stability with respect to different train/test splits. Our algorithm achieves an overall localization error of ~ 2 cm for median. Eyes show the largest localization uncertainty across different runs (largest std), probably due to their smaller size. Furthermore, the use of larger training sets together with global position and shape priors [16] promises to improve generalization across different individuals and anatomies (e.g. missing organs etc.), and increase both the accuracy and its confidence further.

3.2 Comparisons with other algorithms

Gaussian Mixture Models. For comparison we implemented a GMM-based technique where each organ is modelled by fitting a Gaussian Mixture Model to its distribution of CT intensities. During testing, for each voxel $\mathbf{x}$ we evaluate the probability of that point being the centre of a certain organ c. The centre position is then estimated as in (2). Localization errors are reported below:

organ	`head`	`heart`	`l. eye`	`r. eye`	`l. kidney`	`r. kidney`	`l. lung`	`r. lung`	`liver`	*mean acr. organs*
median	53.48	88.54	81.56	85.59	133.04	123.38	89.32	89.59	99.63	93.79 mm
mean	144.27	98.32	174.56	168.42	125.55	128.04	104.88	100.29	98.06	126.93 mm
std	95.63	9.55	121.46	114.19	18.70	15.10	8.48	6.19	14.76	44.90 mm

The table above shows much larger errors than with our technique. An analysis of the posteriors shows that some organ labels are almost uniformly scattered spatially. This induces a bias of the detected centres towards the centre of the volume (thus incorrect), with at times low variance. The reason for such unsatisfactory results is that the GMM approach is based solely on the organs global appearance and fails to capture spatial context; and ways of integrating spatial context *efficiently* within a GMM-based approach are not straightforward. In this case the use of further features such as gradients did not seem to help much.

Template matching. We also compared our technique with a template based method. Here, each organ is represented by a set of 3D templates, extracted from the training volumes and each containing the whole organ. During testing, for each organ c we convolve the input volume with *all* exemplars for that organ and select as centre the point associated with the maximum correlation score over all exemplar templates. Localization errors are reported below.

organ	`head`	`heart`	`l. eye`	`r. eye`	`l. kidney`	`r. kidney`	`l. lung`	`r. lung`	`liver`	*mean acr. organs*
median	167.53	226.54	96.00	98.53	215.31	343.64	230.12	30.89	96.18	167.19 mm
mean	240.31	191.94	238.12	300.05	229.23	303.29	177.18	134.42	150.46	218.33 mm
std	209.08	24.16	51.33	57.20	33.09	67.33	24.41	40.03	55.15	62.42 mm

In this case the results are still worse than with our technique. We believe this is because rigid templates fail to model variations in object's shape, scale and cropping. In this case too the use of gradient features did not help. Finally, as the number of organs of interest increases having to store exemplar templates becomes prohibitive, and the processing burden shifts from training to test.

3.3 Computational efficiency

Training our decision forest model on ~ 26 datasets currently takes around 10 hours on an 8-core Intel desktop. We are planning to port the algorithm onto a High Performance Computing cluster which should reduce training to only about 1 hour. Testing is much faster. In fact, a GPU implementation (following [22]) runs in ~ 2 sec for an approximately 512^3 volume.

4 Conclusion

This paper has introduced a new algorithm for the efficient detection and localization of anatomical structures within Computed Tomography volumes.

We have presented efficient 3D visual features which capture long-range spatial context and help discrimination accuracy. Those features have been incorporated within a random decision forest classifier. The algorithm's parallel nature and the efficiency of its visual features account for the high computational efficiency. The learned model of context accounts for the good localization accuracy.

Next, we plan to extend our technique to other imaging modalities such as MRI, PET-CT and ultrasound. Also, adapting our algorithm to perform hierarchical detection (e.g. thorax $\rightarrow$ heart $\rightarrow$ mitral valve) will help dealing with detailed anatomical structures and will yield richer semantic parsing of medical images. Finally, we would like to extend our work to producing pixel-wise segmentation of complex anatomical structures such as elongated blood vessels. This will necessitate building pixel-wise annotated ground-truth databases and promises to deliver useful results.

References

1. Rubin, G.D.: Data explosion: the challenge of multidetector-row CT. European Journal of Radiology **36**(2) (2000) 74 – 80
2. Linguraru, M.G., Summers, R.M.: Multi-organ automatic segmentation in 4D contrast-enhanced abdominal CT. In: IEEE Intl. Symp. Biom. Im. (ISBI). (2008)
3. Kurkure, U., Avila-Montes, O.C., Kakadiaris, I.A.: Automated segmentation of thoracic aorta in non-contrast CT images. In: IEEE Intl. Symp. Biomedical Imaging (ISBI). (2008)
4. van Ginneken, B., Baggerman, W., van Rikxoort, E.M.: Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans. In: MICCAI. (2008)
5. Shotton, J., Winn, J., Rother, C., Criminisi, A.: Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. In: IJCV. (2009)
6. Shimizu, A., Ohno, R., Ikegami, T., Kobatake, H.: Multi-organ segmentation in three-dimensional abdominal CT images. Int. J CARS **1** (2006)
7. Yao, C., Wada, T., Shimizu, A., Kobatake, H., Nawano, S.: Simultaneous location detection of multi-organ by atlas-guided eigen-orgnmethod in volumetric medical images. Int. J CARS **1** (2006)
8. Han, X., Hoogeman, M.S., Levendag, P.C., Hibbard, L.S., Teguh, D.N., Voet, P., Cowen, A.C., Wolf, T.K.: Atlas-based auto-segmentation of head and neck CT images. In: MICCAI. (2008)
9. Zhuang, X., Rhode, K., Arridge, S., Razavi, R., Hill, D., Hawkes, D., Ourselin, S.: An atlas-based segmentation propagation framework using locally affine registration – application to automatic whole heart segmentation. In: MICCAI. (2008)
10. Fenchel, M., Thesen, S., Schilling, A.: Automatic labeling of anatomical structures in MR fastview images using a statistical atlas. In: MICCAI. (2008)
11. Dolejst, M., Kybic, J., Tuma, S., Polovincak, M.: Reducing false positive responses in lung nodule detector system by asymmetric adaboost. In: ISBI. (2008)

12. Pescia, D., Paragios, N., Chemouny, S.: Automatic detection of liver tumors. In: ISBI. (2008)
13. Wels, M., Carneiro, G., Aplas, A., Huber, M., Hornegger, J., Comaniciu, D.: A discriminative model-constrained graph-cuts approach to fully automated pediatric brain tumor segmentation in 3D MRI. In: MICCAI. (2008)
14. Freiman, M., Edrei, Y., Shmidmayer, Y., Gross, E., Joskowicz, L., Abramovitch, R.: Classification of liver metastases using fMRI images: A machine learning approach. In: MICCAI. (2008)
15. Prasad, M., Sowmya, A.: Multi-level classification of emphysema in HRCT lung images using delegated classifiers. In: MICCAI. (2008)
16. Tu, Z., Narr, K.L., Dollar, P., Dinov, I., Thompson, P.M., Toga, A.W.: Brain anatomical structure segmentation by hybrid discriminative/generaltive models. IEEE Trans. on Medical Imaging **27**(4) (2008)
17. Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Toga, A.W., Thompson, P.M.: Automatic subcortical segmentation using a contextual model. In: MICCAI. (2008)
18. Pohl, K.M., Bouix, S., Nakamura, M., Rohlfing, T., McCarley, R.W., Kikinis, R., Grimson, W.E.L., Shenton, M.E., Wells, W.M.: A hierarchical algorithm for MR brain image parcelation. IEEE Trans. on Medical Imaging **26**(9) (2007)
19. Zhan, Y., Zhou, X.S., Peng, Z., Krishnan, A.: Active scheduling of organ detection and segmentation in whole-body medical images. In: MICCAI. (2008)
20. Yin, P., Criminisi, A., Essa, I., Winn, J.: Tree-based classifiers for bilayer video segmentation. In: CVPR. (2007)
21. Bosch, A., Zisserman, A., Munoz, X.: Image classification using random forests and ferns. In: IEEE ICCV. (2007)
22. Sharp, T.: Implementing decision trees and forests on a GPU. In: ECCV. (2008)
23. Breiman, L.: Random forests. Technical Report TR567, UC Berkeley (1999)
24. Amit, Y., Geman, D.: Shape quantization and recognition with randomized trees. Neural Computation **9** (1997) 1545–1588
25. Quinlan, J.R.: C4.5: Programs for Machine Learning. (1993)
26. Viola, P., Jones, M.J.: Robust real-time face detection. IJCV (2004)
27. Dalal, N., T.B.: Histograms of oriented gradients for human detection. In: IEEE CVPR. (2005)
28. Zambal, S., Buehler, K., Hladuvka, J.: Entropy-optimized texture models. In: MICCAI. (2008)
29. Crow, F.C.: Summed-area tables for texture mapping. In: SIGGRAPH '84: Proceedings of the 11th annual conference on Computer graphics and interactive techniques, New York, NY, USA, ACM (1984)
30. Viola, P., Jones, M.J., Snow, D.: Detecting pedestrians using patterns of motion and appearance. In: ICCV. (2003)
31. Torralba, A., Murphy, K.P., Freeman, W.T.: Sharing visual features for multiclass and multiview object detection. IEEE Trans. PAMI (2007)