A Generic Approach to Organ Detection Using 3D Haar-Like Features

Florian Jung[1], Matthias Kirschner[2], Stefan Wesarg[1]

[1]Cognitive Computing & Medical Imaging, Fraunhofer IGD, Darmstadt, Germany
[2]Graphisch-Interaktive Systeme, TU Darmstadt, Darmstadt, Germany
florian.jung@igd.fraunhofer.de

Abstract. Automatic segmentation of medical images requires accurate detection of the desired organ as a first step. In contrast to application-specific approaches, learning-based object detection algorithms are easily adaptable to new applications. We present a learning-based object detection approach based on the Viola-Jones algorithm. We propose several extensions to the original approach, including a new 3D feature type and a multi-organ detection scheme. The algorithm is used to detect six different organs in CT scans as well as the prostate in MRI data. Our evaluation shows that the algorithm provides fast and reliable detection results in all cases.

1 Introduction

The segmentation of organs is one of the major tasks in medical imaging. Common segmentation algorithms operate locally, that is why a preceding detection of the organ is needed to approximate its position. Many segmentation algorithms still need a manual placement of a seed point in order to successfully segment the desired structure. In order to achieve a fully automatic organ segmentation, the detection task has to be automated as well. Many object detection methods are tailored to a specific application. As an example, Kainmüller et al. [1] detect the liver by searching for the right lung lobe, which can be relatively easily detected by thresholding and voxel counting. The problem of such approaches is that they do not generalize to other structures or image modalities. Learning-based approaches such as Discriminative Generalized Hough transform [2] and Marginal Space Learning (MSL) [3, 4] are more general and can be adapted to a wide variety of detection tasks by simply exchanging the training data.

In this paper, we propose a learning-based organ detection approach for 3D medical images based on the Viola-Jones object detection algorithm [5]. We use a bootstrapping approach to automatically select important training data. Moreover, we propose several extensions to the original approach tailored to medical images, namely a new pure-intensity feature type, and a multi-organ detection that exploits spatial coherence in medical data. Finally, we provide a detailed evaluation of the algorithm by detecting six different organs in CT data sets as well as the prostate in T2-weighted MRI scans. Hereby, we show that our algorithm can successfully detect these seven different structures.

H.-P. Meinzer et al. (Hrsg.), *Bildverarbeitung für die Medizin 2013*, Informatik aktuell,
DOI: 10.1007/978-3-642-36480-8_56,

2 Methods

The basis of this algorithm is the Viola-Jones face detection [5] which builds a strong classifier based on Haar-like features. We use AdaBoost to train this classifier and to select a subset of 3D Haar-like features [4] within an iterative training process. We made several modifications to the algorithm in order to be able to use it in 3D.

2.1 Bootstrapped learning process

The number of subregions per image is huge and the time needed to train a classifier using all the regions is not negligible. Therefore we extract only a subset of subregions from the training images. In order to extract regions with high information value for the training process we implemented a bootstrapped process which trains a classifier and then starts a detection on the training images. All resulting false positives are added as additional negative samples to the already existing examples. Finally a new training process is executed. With the use of these additional samples, we expect that the classifier can better discriminate regions that are hard to classify.

2.2 Adjusting the ROI

In most cases, we use the organ's bounding box as positive training example. As the tissue within an organ is relatively homogeneous and 3D Haar features describe the intensity contrast between adjacent regions, 3D Haar features are usually more discriminative when they cover tissue of both the organ and its surroundings. In cases where the bounding box alone does not provide sufficiently discriminative features, the ROI can be enlarged by a constant factor to include additional tissue. After detection, the detected bounding box can be easily reconstructed by downsizing the enlarged ROI.

2.3 New feature type

A Haar-like feature consists of two (or more) rectangular areas which are summed up and subtracted from each other. The difference between these area-sums is the resulting Haar-like feature value. A Haar-like feature is defined by its type, its size and its position inside the current window.

In order to account for regions with approximately constant image intensities that are often observed in CT images we create a new feature type that only sums up all pixels within its area without doing any subtraction. Therefore we expect the feature we added to be an enrichment for our algorithm and improve the detection results.

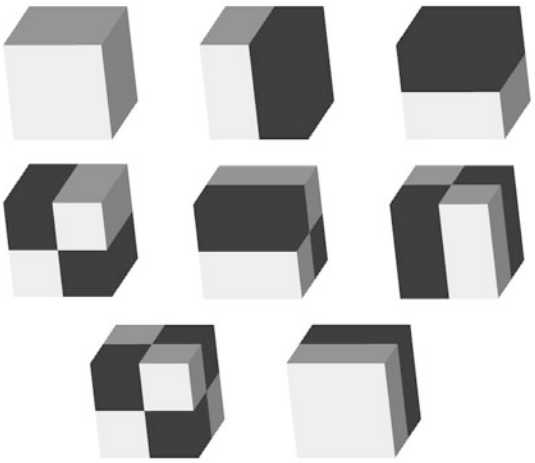

Fig. 1. Illustration of the eight 3D Haar-like feature types we are using for our organ detection algorithm.

2.4 Detection

The organ detection is done using the sliding window. At each window position all feature values for the selected subset of 3D Haar-like features are calculated, multiplied with the weight that was stored during the training process and finally summed up. If this value exceeds the threshold defined during the training phase, it is assumed that the region contains the relevant structure. The detection is done using 5 different sliding window sizes. Starting with the mean size of the organ the window is downsized twice and upsized twice according to observed organ size variations. In contrast to face detection, where every image can potentially contain many faces, or no face at all, we can safely assume that there is exactly one organ per image. Therefore, the final detection result is the sub window with the highest signed distance to the decision threshold.

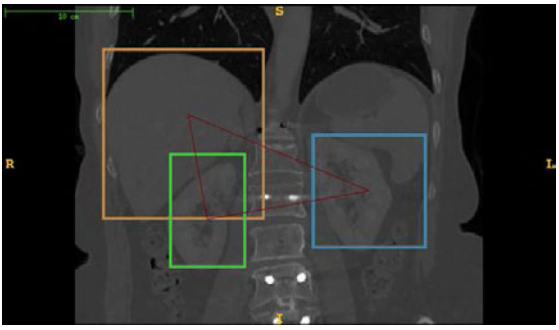

Fig. 2. Illustration of the relationship model between different organs. The usage of this model improves the detection results significantly.

2.5 Multi-organ detection

In order to further improve the detection results and to be able to detect organs which could not be detected with the default detection approach, we implemented a multi-organ detection. For this purpose, we use a PCA (principal component analysis) and train a model consisting of the organs of interest and their relative positions to each other using the center of gravity. During the detection phase we analyze the positions of the detected organs and compare them with the trained PCA model. A probability function returns us the likelihood of a valid organ model. We create a ranking of the 30 best results for each organ and than calculate a detection score for each possible permutation considering the probability as an additional weight.

2.6 Experimental setup

For tests on contrast enhanced CT data, we used 210 images to train the classifier and 10 images to evaluate the results. The images contained the following organs: liver, heart, bladder, kidneys and the spleen. The spacing of the images varies between: x: 0.57, y: 0.57, z: 1 and x: 0.97, y: 0.97, z: 5. The detected ROIs are compared with the ground truth bounding boxes of the images using the Dice Coefficient. We trained more than 250 classifiers for different organs and with varying window size. Performing some tests we found out that an Dice overlap of 60-70 % for the detection is enough to accomplish a good segmentation.

For tests on MR data, we detected the prostate on T2-weighted MRI scans. The data was obtained from the MICCAI Grand Challenge: Prostate MR Image Segmentation 2012. For training the detector, we used the 50 training images, and then performed tests on the 30 test images. As the ground truth of the 30 test scans is not available to us, the results were assessed by visual inspection only. Because in contrast to CT images, MRI data has non-standardized image intensities, all images were pre-processed. Intensity inhomogeneities within a single a image, the so-called bias-field, was removed with Coherent Local Intensity Clustering [6]). Afterwards, image intensities are normalized using robust statistics and rescaled to the interval $[0, 1000]$.

Adjusting the ROI on MRI. Because the tissue within the prostate's bounding box is very inhomogeneous, we adjusted the ROI around the prostate (see Section 2.2) in x and y-direction by 40 % in order to achieve more robust detection results.

Multi-organ detection. The multi-organ detection was applied to the task of simultaneously detecting liver and both kidneys, because the latter are difficult to detect. As test images we choose images in which the detection of at least one organ failed.

3 Results

Fig. 3 shows quantitative results obtained on the CT data with a detection window size of 10^3. We achieve a median Dice coefficient with the ground truth bounding box of 0.71-0.87, depending on the data set. Only for the kidneys, we observed misdetections.

The size of the sliding window has major impact on the training time. A training run with a resolution of 5^3 can be accomplished within 5 minutes, while a classifier trained with a resolution of 15^3 needs two and a half days. However, we observed that the classifiers trained with a resolution of 5^3 already yield some usable detection results. Here, the median Dice coefficient is between 0.62 and 0.78.

On 30 MRI test scans, visual inspection showed that the detection results were sufficient for segmentation. Indeed, in a subsequent Active Shape Model segmentation, we achieved a median Dice coefficient of 0.86 [7].

The detection itself is very fast. Although it depends on the image size, the size of the sliding window and some other factors, it takes less than one second in the majority of cases.

3.1 New feature type

The new feature type we created is used in 15 per cent of cases for an average classifier and was thereby the second most used classifier. Considering that the boosting algorithm chooses the feature that can distinguish best between the samples, it can be said that the inclusion of the new feature type turns out to be an enrichment for the algorithm.

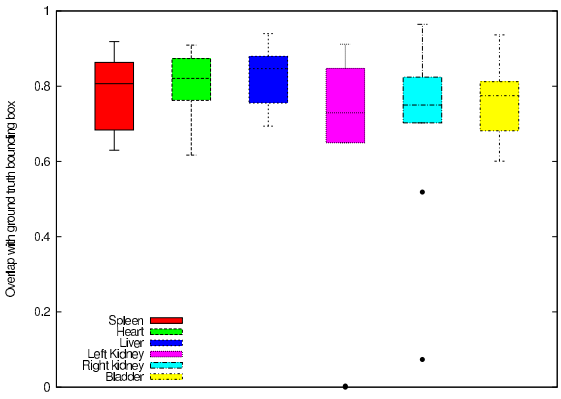

Fig. 3. Illustration of the overlaps of six different organs. Comparison was done between the ground-truth bounding-box and the detected bounding box. All results are located between 0.6 and 0.95 %(except for three kidney results), what turned out to be enough to do a proper segmentation of the organ.

Multi-organ detection. The use of the multi-organ detection improves the detection results for organs that are more difficult to detect (like the kidneys)

significantly. In 75% of the cases the missclassified organ was detected due to the usage of the multi-organ detection. Additionally the average overlap of the three detected organs increased by 20% thanks to this approach.

4 Discussion

We proposed several extensions to the Viola-Jones face detection algorithm and applied them to the detection of organs in 3D medical images. Our approach achieves good detection results on seven different organs, including six CT and one MRI data set. This shows the broad applicability of our approach in contrast to methods tailored to specific applications. In cases where the detection is difficult, for example for the kidneys, the use of the multi-organ detection improves the results and enables use to locate organs that can not be robustly detected with a single detector. Alternatively, it might be possible to improve the results by enlarging the ROI, as it was done for the prostate.

In contrast to MSL, we do not estimate the orientation of the organ. This estimation is, in most cases, not necessary because of the scanning protocol. For example, patients are always oriented upright in CT images. We also use a simple learning algorithm, namely boosting instead of the probabilistic boosting tree used in MSL. However, our results show that this simpler algorithm provides sufficient results in practice, and is applicable to a large variety of different organs.

References

1. Kainmüller D, Lange T, Lamecker H. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. Proc MICCAI Grand Challenge: 3D Segmentation in the Clinic. 2007; p. 109–16.
2. Ruppertshofen H, Lorenz C, Schmidt S, et al. Discriminative generalized hough transform for localization of joints in the lower extremities. CSRD. 2011;26:97–105.
3. Ling H, Zhou SK, Zheng Y, et al. Hierarchical, learning-based automatic liver segmentation. Proc IEEE CVPR. 2008; p. 1–8.
4. Zheng Y, Barbu A, Georgescu B, et al. Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. IEEE Trans Med Imaging. 2008;27(11):1668–81.
5. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. Proc CVPR. 2001;1:511–18.
6. Li C, Xu C, Anderson A, et al. MRI tissue classification and bias field estimation based on coherent local intensity clustering: A unified energy minimization framework. IPMI. 2009; p. 288–99.
7. Kirschner M, Jung F, Wesarg S. Automatic prostate segmentation in MR images with a probabilistic active shape model. Proc MICCAI Grand Challenge: Prostate MR Image Segmentation. 2012.