

# Regression Forest-Based Organ Detection in Normalized PET Images

Peter Fischer<sup>1</sup>, Volker Daum<sup>2</sup>, Dieter Hahn<sup>2</sup>, Marcus Prümmer<sup>2</sup>,  
Joachim Hornegger<sup>1</sup>

<sup>1</sup> Pattern Recognition Lab and Erlangen Graduate School in Advanced Optical  
Technologies (SAOT), FAU Erlangen-Nürnberg

<sup>2</sup> Chimaera GmbH

`peter.fischer@fau.de`

**Abstract.** The detection of organs from full-body PET images is a challenging task due to the high noise and the limited amount of anatomical information of PET imaging. The knowledge of organ locations can support many clinical applications like image registration or tumor detection. This paper is the first to propose an organ localization framework tailored on the challenges of PET. The algorithm involves intensity normalization, feature extraction and regression forests. Linear and nonlinear intensity normalization methods are compared theoretically and experimentally. From the normalized images, long-range spatial context visual features are extracted. A regression forest predicts the organ bounding boxes. Experiments show that percentile normalization is the best preprocessing method. The algorithm is evaluated on 25 clinical images with a spatial resolution of 5 mm. With 13.8 mm mean absolute bounding box error, it achieves state-of-the-art results.

## 1 Introduction

In order to correctly use Positron Emission Tomography (PET) images for diagnosis, it is helpful for physicians to know the relation to the underlying anatomy. This information can be provided with the help of morphological images, for example Computed Tomography (CT) or Magnetic Resonance (MR) images. Nowadays, the standard scanners are hybrids of functional and morphological modalities, e.g. PET/CT or PET/MR. Even in hybrid scanners, the images are not perfectly aligned due to motion artifacts and different acquisition times. This could be tackled by image registration. Image registration can be greatly enhanced by semantic information like known organ positions. Most of the publications on anatomy localization deal with CT images.

Seifert et al. segment organs and detect point landmarks automatically from full body CT images [1] using marginal space learning with probabilistic boosting trees and 3-D Haar features. Detection of landmarks and organs is also performed in [2] with classification forests. The features used by Criminisi et al. are a generalization of 3-D Haar features called visual features which emphasize long-range spatial context. This work was extended in [3] by using regression forests instead

of classification forests. Pauly et al. transferred the approach using regression ferns and binary visual features to MR imaging [4]. In PET imaging literature, the focus of automatic localization methods is not on anatomy, but on tumors. Guan et al. include a rough body part localization by classifying feature curves from the PET volume using a hidden Markov model [5]. Montgomery et al. report a fully automated, unsupervised segmentation of PET volumes using Gaussian mixture models and a multiscale Markov random field [6].

This work localizes organs in the challenging environment of PET imaging using context-rich visual features, regression forests [3,4], and intensity normalization. Multiple intensity normalization methods are analyzed and compared experimentally.

## 2 Materials and Methods

### 2.1 Intensity Normalization

Attenuation corrected PET images contain measurements of the count of positron emission decays in each voxel. The counts are a physical quantity, but they are not directly comparable for different acquisitions as Hounsfield units in CT. Reasons are the variability caused by the scanner, the injected dose, the tracer, the uptake time, and the human anatomy. Intensity normalization reduces the variability that the organ detection algorithm has to deal with.

The normalization of *mean and variance* of an image  $I$  makes subsequent features invariant to affine changes of intensity. An affine change of intensity is  $I'(\mathbf{v}) = a I(\mathbf{v}) + b$  with the voxel  $\mathbf{v} \in \mathbb{R}^3$  and constants  $a, b \in \mathbb{R}$ . In normalization, the mean  $\mu$  is set to 0, and the variance  $\sigma^2$  to 1 using the transform  $I_{MV}(\mathbf{v}) = \frac{I(\mathbf{v}) - \mu}{\sigma}$ . This nullifies the effects of affine intensity changes. *Percentile* normalization is in principle the same as the normalization of minimum and maximum values of the image. The difference is that outliers in the image are saturated before normalization. For the non-outlier intensities, this normalization removes affine transformations of the intensity. The percentile normalization works by  $I_{PERC}(\mathbf{v}) = \frac{I(\mathbf{v}) - I_{low}}{I_{high} - I_{low}}$ , with the low percentile  $I_{low}$  and the high percentile  $I_{high}$  set to intensities corresponding to an arbitrary percentage of the intensities of all voxels in the image. *Standard uptake values* (SUV) are used in radiology to diagnose the malignancy of tumors in PET imaging. SUV is the normalization  $I_{SUV}(\mathbf{v}) = \frac{I(\mathbf{v})}{D/BW}$ , with the patient body weight  $BW$  [g], the injected dose  $D$  [Bq], and the radioactivity concentration in a given voxel  $I(\mathbf{v})$  [Bq/ml] [7]. The major difference between the above affine normalization schemes is the information source. Mean and variance normalization uses moments, whereas percentile normalization uses quantiles of the histogram. Both methods use only image information, as opposed to SUV which uses clinical and patient meta-information. The true transformation between images, especially of different patients, is not affine. Consequently, the normalizations presented so far cannot remove these variations. Nonlinear intensity variations also occur in magnetic resonance imaging. We transfer a sophisticated normalization scheme

in MR to PET, namely *non-rigid registration* of the image histogram to a reference histogram [8]. To avoid that the registration algorithm focuses on the background intensities, which constitute a dominant peak in the histogram, and to limit the number of histogram bins, extreme intensities are excluded from the registration [8]. This is achieved by masking values below the 80 % percentile and above the 99 % percentile.

## 2.2 Organ Localization

Due to the human anatomy and the characteristics of imaging systems, the appearance and the relative position of anatomical structures is similar in medical images of the same modality. Consequently, the high amount of contextual information that is available should be exploited by the features. A successful feature framework in medical object localization are the *visual features* from Criminisi et al., which capture anatomical structures and their relative positions [2,3,4]. The features consist of the relationship of cuboid regions with a random offset to the voxel under consideration and with random size. One benefit of visual features is that context can be captured well over a long range. Visual features are especially suited for PET imaging because the involved averaging makes them robust to the high noise of PET and the context information allows them to cope with the low discriminability of small voxel neighborhoods. Instead of intensity normalization, it is possible to make the features themselves invariant to intensity changes by replacing the difference in [2] with a binary comparison [4]. However, information about the relative magnitude of the features is lost because of the binarization. Another issue is that some feature regions necessarily lie outside of the image for some voxels due to the random offsets. For full-body PET images, the feature value for these boxes can set to zero, assuming no tracer concentration outside of the image. This assumption is only violated in the area of the legs, but as the field of view is similar for all images, the violation has only a small influence.

In this work, organs in PET images are detected using the nonlinear, multidimensional regression algorithm of *regression forests*. The output of the regression is a vector containing the bounding boxes of several organs [3]. The cost function for tree training is based on class affiliation [9], which in this case are the organs. The affiliation to an organ is modeled by a Laplacian density with the empirically set parameter  $\lambda = 100$  mm. The number of features that are examined during training in each node is called randomness  $\rho$ . In each leaf of the tree, a Gaussian density is stored as a probabilistic approximation of the training samples. During testing, the trees of the forest are combined by weighted averaging. For an image, all the samples are combined by adding the mean value of the Gaussian of the leaf in which the samples end up in, weighted by the inverse of the trace of the covariance matrix of the respective leaf. In addition, only the 10 % of samples with highest weight are retained, the others are discarded. Regression forests are well suited because they incorporate multidimensional outputs, are fast, and easy to parallelize. Regression is superior to voxel-wise classification in

PET imaging, because the involved averaging makes the estimates robust to the high noise level.

### 2.3 Experiments

**Comparison of Intensity Normalization Methods.** The intensity normalization methods are compared using organ localization and histogram errors. The organ localization error is the absolute difference of the estimated and the true bounding boxes. Measures to compare the PET image histograms are the Sum of Squared Differences (SSD) and the symmetric Kullback-Leibler Divergence (SKLD). The histograms are computed with 256 equally sized bins, except for percentile normalization, where the number of bins is reduced to reflect the compression of the intensity range. The histogram error reflects the similarity of the intensity distributions of different PET images.

The random forest in this experiment consists of 8 trees with a maximal depth of 10.  $\rho = 100$  visual features are searched in each node during forest training with a maximum size of 200 mm and offset of 500 mm. Training is performed on 16 images and testing on 4 images not in the training set. In percentile normalization, 0% and 99.5% percentiles are used.

**Organ Localization.** In the PET images, brain, liver, left and right kidney, and bladder are located. The regression forest has 14 trees of depth up to 9. The randomness is  $\rho = 500$ . Percentile normalization is performed with 0% and 99.5%. The visual features are thresholded to a maximum size and offset of 200 mm. There are 25 PET images with a spatial resolution of 5 mm in each direction. All are acquired with the tracer fluorodeoxyglucose (FDG). From each image, 1% of the voxels are drawn for training to reduce training time. The results are computed using 5-fold cross validation.

## 3 Results

### 3.1 Comparison of Intensity Normalization Methods.

The results are summarized in Tab. 1. The lowest histogram errors are achieved by SUV normalization. This shows that SUV is a good choice in clinical evaluation of PET images. Percentile normalization reduces the SKLD error, but shows no improvement in SSD. Non-rigid histogram registration is most flexible and thus expected to give the lowest errors. Instead, it increases both errors, because PET histograms show little structure that can be aligned [5]. As histogram registration is worsening the total histogram alignment, its organ localization error is not investigated. The organ localization test errors agree only partially with the histogram errors. Without preprocessing, the algorithm achieves a test error of 15.6 mm. Mean-variance normalization and binary comparison increase the test and train error, which can be explained by the change of the background intensity and by the missing magnitude information, respectively. With 11.7 mm

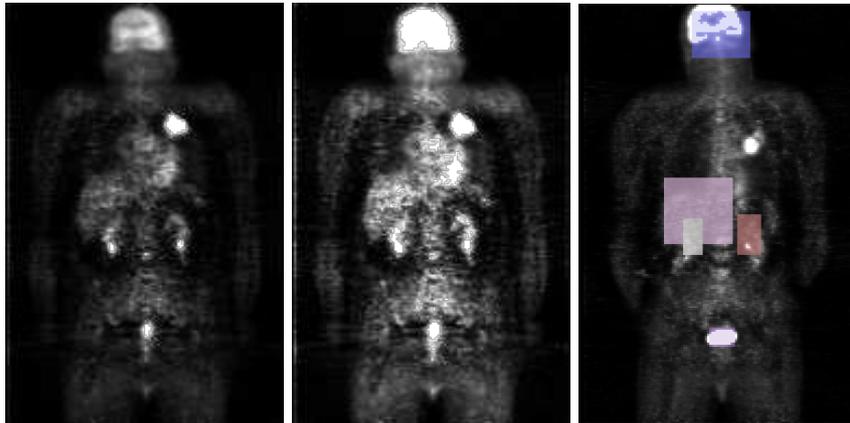
**Table 1.** This table compares mean absolute bounding box errors over all organs and the histogram errors for different preprocessing methods.

Preprocessing Method	Train	Test	Histogram	Histogram
	Error [mm]	Error [mm]	Error [SKLD]	Error [SSD]
None	8.1	15.6	0.055	$3.1e12$
Mean-Variance	28.5	26.9	0.034	$6.3e12$
Percentile	6.5	11.7	0.013	$3.5e12$
SUV	6.7	13.9	0.0004	$1.6e10$
Histogram Registration	–	–	0.061	$3.4e12$
Binary Comparison	9.7	18.1	–	–

test error, percentile normalization is the best in our experiments. The left and middle of Fig. 1 show an unprocessed and a percentile normalized PET image.

### 3.2 Organ Localization.

The mean organ localization error is  $13.8 \pm 7.5$  mm. This is higher than the corresponding experiment in Tab. 1 due to the cross-validation. Separated into single organs, we achieve an error of  $15.4 \pm 11.3$  mm for the brain,  $13.0 \pm 4.9$  mm for the liver,  $13.4 \pm 6.5$  mm for the left and  $11.1 \pm 5.5$  mm for the right kidney, and  $15.9 \pm 6.4$  mm for the bladder. The errors compare favorably with the ones from the literature, e.g. 15.0 mm for organ localization in MR [4] and 16.7 mm for organ localization in CT [3]. The average runtime of training is 5 h for 20 images and the runtime of testing is 10 min for 5 images. In Fig. 1, a typical organ localization result is shown overlaid onto the PET image.



**Fig. 1.** *Left* is a slice of an attenuation-corrected PET image. In the *middle* is the same slice with normalized percentiles. On the *right*, a typical organ localization result is shown.

## 4 Discussion

This paper presents an algorithm for automatic detection of organs in PET images. This is achieved with nonlinear regression of organ bounding boxes using a regression forest. Visual features with long-range spatial context are the input of the regression algorithm. Different intensity normalization methods are proposed to cope with the unstandardized intensities of PET. Experimental evaluation shows that percentile normalization works best because it removes outliers and affine transformations between the image intensities. The final organ localization error, computed using 5-fold cross-validation, is 13.8 mm on average over all images and organs. In the future, the inclusion of organs and landmarks into image registration could be analyzed. It should be determined whether any improvement over traditional PET/CT and PET/MR is achievable.

### Acknowledgments

We thank the nuclear medicine department of the Universitätsklinikum Erlangen for providing the images and medical advice. The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative.

### References

1. Seifert S, Barbu A, Zhou SK, et al. Hierarchical Parsing and Semantic Navigation of Full Body CT Data. *Proc SPIE*. 2009;7259:725902–725902–8.
2. Criminisi A, Shotton J, Bucciare S. Decision Forests with Long-Range Spatial Context for Organ Localization in CT Volumes. In: Wells III WM, et al, editors. *Proc MICCAI PMMIA*. Springer; 2009. p. 69–80.
3. Criminisi A, Shotton J, Robertson DP, et al. Regression Forests for Efficient Anatomy Detection and Localization in CT Studies. In: Menze BH, et al, editors. *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. Springer; 2010. p. 106–117.
4. Pauly O, Glocker B, Criminisi A, et al. Fast Multiple Organ Detection and Localization in Whole-Body MR Dixon Sequences. In: Fichtinger G, et al, editors. *Proc MICCAI, Part III*. Springer; 2011. p. 239–247.
5. Guan H, Kubota T, Huang X, et al. Automatic Hot Spot Detection and Segmentation in Whole Body FDG-PET Images. In: *Proc ICIP*. IEEE; 2006. p. 85–88.
6. Montgomery DWG, Amira A, Zaidi H. Fully Automated Segmentation of Oncological PET Volumes using a Combined Multiscale and Statistical Model. *Med Phys*. 2007;34:722–736.
7. Kelly M. SUV: Advancing Comparability and Accuracy. Siemens; 2009.
8. Jäger F, Hornegger J. Nonrigid Registration of Joint Histograms for Intensity Standardization in Magnetic Resonance Imaging. *IEEE Trans Med Imaging*. 2009;28(1):137–150.
9. Dantone M, Gall J, Fanelli G, Van Gool L. Real-Time Facial Feature Detection using Conditional Regression Forests. In: *Proc CVPR*. IEEE; 2012. p. 2578–2585.