

# Automatic Recognition of Anatomical Regions in Computed Tomography Images

Márton József Tóth<sup>1\*</sup>, László Ruskó<sup>2</sup>, Balázs Csébfalvi<sup>1</sup>

<sup>1</sup> Department of Control Engineering and Information Technology, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, H-1117 Budapest, Magyar tudósok körútja 2, Hungary

<sup>2</sup> GE Hungary Healthcare Division, H-6722 Szeged, Petőfi Sándor sgt 10, Hungary

\* Corresponding author, e-mail: [tmarton@iit.bme.hu](mailto:tmarton@iit.bme.hu)

Received: 23 July 2018, Accepted: 16 October 2018, Published online: 04 December 2018

## Abstract

This paper presents a method that can recognize anatomy regions in Computed Tomography (CT) examinations. In this work the human body is divided into eleven regions from the foot to the head. The proposed method consists of two main parts. In the first step, a Convolutional Neural Network (CNN) is used to classify the axial slices of the CT exam. The accuracy of the initial classification is 93.4 %. As the neural network processes the axial slices independently from each other, no spatial coherence is guaranteed. To ensure the contentious labeling the initial classification step is followed by a post-processing method that incorporates the expected order and size of the anatomical regions to improve the labeling. In this way, the accuracy is increased to 94.0 %, the confusion of non-neighboring regions dropped from 1.5 % to 0.0 %. This means that a continuous and outlier free labeling is obtained. The method was trained on a set of 320 CT exams and evaluated on another set of 160 cases.

## Keywords

anatomy region recognition, deep learning, image classification, imaging informatics, medical image processing

## 1 Introduction

Three-dimensional medical imaging techniques are routinely applied in the clinical practice. Computed Tomography (CT) as well as Magnetic Resonance Imaging (MRI) are integral parts of diagnosis, therapy planning, and monitoring. As the number of the medical images is continuously increasing [1], the computer aided processing of these images becomes more and more important.

To provide support for the therapists several algorithms have been developed for various clinical workflows, like segmentation, registration or visualization. These methods are usually specialized to a specific region of interest (ROI). There are different applications for lung nodule detection [2], virtual colonoscopy [3], cardiac or cerebral vessel analysis [4], tumor follow-up [5, 6] etc.

The DICOM standard involves tags to specify the anatomy location for each slice of an image series, but it is rarely filled in by the scanner and the manual specification is time consuming and error prone [7]. Automated pre-processing functions cannot rely on this information, which can lead to incomplete processing of

medical images that requires time consuming correction from the user.

From the above-mentioned examples one can see that the automated detection of anatomical regions in medical images would have great impact to the content-based medical image processing. It would create many opportunities to automate or optimize various types of algorithms (i.e. initialization of segmentation or registration methods), which would save significant amount of time and workload for the clinicians. Although the need for this function does not directly come from the user, its benefits are definitely welcome in medical image processing.

In this work, we present a method that can reliably recognize the main anatomical regions in 3D CT images. The recognition process starts with an image-preprocessing step and a preliminary recognition process that utilizes convolutional neural networks. These standard steps are followed by a newly introduced post-processing method that can filter the initial classification results to ensure a correct final labeling.

## 2 Related Work

Over the last few years, deep learning methods became popular in the computer vision community. The convolutional neural networks (CNN) are especially designed to work with image processing tasks (object detection, classification or segmentation), as images contain highly correlated local intensity patterns. The CNN methods [8–10] regularly outperform the traditional image feature extraction based methods in image classification tasks.

As the automatic classification of human anatomy is an important prerequisite in many computer aided diagnostic processes, many researchers aimed different level of recognition of the anatomical regions. There were approaches to identify anatomic landmark points [11–13]. These methods provided good results in case of full body CT scans where all the required landmarks are visible.

Other techniques were developed to classify the axial slices of scans using a manifold learning method [14, 15]. The reported accuracy is around 94.0% considering six classes.

The recent VISCERAL challenges [1, 16] have also showed that the recognition of body parts is a very active field of research and could be well utilized in clinical practice.

CNNs were also applied for recognition of anatomical regions. Yan et al. [17] provided a great overview about the application of different CNN configurations in the classification of axial images. Roth et al. [18] used five classes to identify axial slices of CT images and achieved the state-of-the-art accuracy (5.9 % error).

In summary, the classification of the axial slices of CT images is widely being researched. The recent CNN classification methods provide state-of-the-art quality results. However, the CNNs use statistical models and therefore can produce noisy output, which cannot be directly used in the clinical practice. It is worth incorporating a priori information of the human body to stabilize the results of the CNN and to provide a reliable classification.

## 3 The Data Set

The CT examinations used in this work (for training as well as evaluation) were originally collected to evaluate various types of clinical applications in the area of oncology, cardiology, surgery, neurology. The cases involve whole-body as well as partial scans. The images were acquired with whole (50-70 cm) as well as small (20-35 cm) reconstruction diameter. The dataset shows significant variance in patient's sex, age, and level of obesity. It includes native and contrast-enhanced cases, and exams demonstrating

various types of abnormality (pathology, implant, noise). Thus, it is a representative set of 480 CT images.

The definition of anatomy regions: The primary goal of this work was to facilitate automated processing of CT images. The definition of the anatomy regions was driven by the needs of organ segmentation algorithms (e.g. brain, lung, liver, colon, prostate). Based on this consideration eleven anatomy regions were defined in agreement with radiologists as defined in Table 1.

Each region starts at the axial slice involving the starting point and ends before the first slice of the next region. The above defined anatomy landmarks were manually defined by radiologists for all examinations, and each slice in the dataset was automatically labeled with the corresponding region label based on its location. Table 2 summarizes the statistical properties of the established regions. This information is later utilized in Section 8.

**Table 1** Region definitions.

Region	Starting Point
FOOT	bottom of foot
SHIN	articular cavity of the ankle
THIGH	articular cavity of the knee
PELVIS-LOWER	separation of legs
PELVIS-CENTER	bottom of symphysis
PELVIS-UPPER	articular cavity of the hip
ABDOMEN-LOWER	top of the iliac crest
ABDOMEN-UPPER	bottom of the lateral sinuses
CHEST	top of the liver
HEAD	cranial end plate of the first thoracic vertebra
BRAIN	foramen magnum

**Table 2** Region size statistics.

Name	Average [mm]	Max [mm]	Min [mm]
FOOT	125	200	23
SHIN	292	284	35
THIGH	260	260	221
PELVIS-LOWER	55	80	13
PELVIS-CENTER	45	64	30
PELVIS-UPPER	141	162	101
ABDOMEN-LOWER	140	210	97
ABDOMEN-UPPER	64	125	40
CHEST	175	252	105
HEAD	119	141	92
BRAIN	91	160	63

#### 4 Method

The proposed method consists of three main steps. In the first step, traditional image processing methods are used to normalize the axial slices of the CT scan. With this step the most interfering artifacts can be eliminated and the slices can be used in a uniform way in the further processing steps.

The next stage of the method is the initial classification of the axial slices. In this step, a CNN is used to assign anatomy labels to the slices. In addition to the most likely label, the confidence for each label can also be obtained. These confidences play important role in the post-processing step.

In the last phase, the initial labeling and the region confidences are incorporated. The output of the CNN can be noisy and can contain classification errors, as it is a result of a statistical model. To obtain a result that is usable in clinical practice, some properties of the final labeling should be ensured. This final step guarantees that the order of the detected anatomical regions is correct and all the detected regions have a realistic size.

#### 5 Image Processing

As the imaging options can vary from scan to scan, a pre-processing step is applied before slice classification.

The position of the patient with respect to image center can vary significantly among examinations, which introduces unwanted variation. Even if the patient is acquired in supine position the table can be lifted in anterior or posterior direction. In order to compensate that, the weight center of non-air voxels is computed that is used to define the body axis for the whole CT exam.

After this point each slice is separately processed. First, a squared region is extracted, such that its center is located on the body center axis (that is defined as the average of the weight center for all slices), its size (both width and height) is equal to 35 cm, and its resolution is equal to 256x256 pixels. Since the pixel size of CT images varies among examinations interpolation is used. If the original input slice covers larger or smaller image region than 35 cm, the pre-processed image is cropped or padded (with air voxels), respectively. Then, the pixel intensities (from the original 16-bit signed integer value) are rescaled to the range [0,255], such that the range [-500,500] HU is linearly mapped to [0,255]. This step enhances those tissue types which are key important from anatomy point of view and prevents the classifier taking very high or low density pixels into account.

Fig. 1 shows the result of the pre-process. In consequence of this step the most interfering imaging options

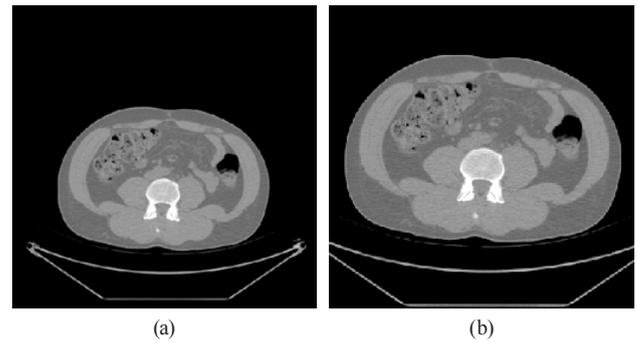


Fig. 1 Pre-process result. (a) Original slice, (b) Pre-processed slice.

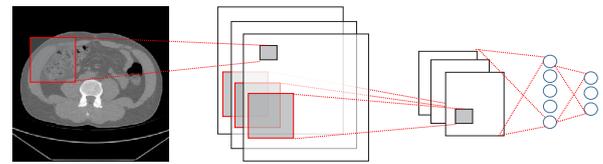


Fig. 2 CNN topology example.

can be normalized and the resulted axial slices can be used in a uniform way in the following steps.

#### 6 Convolutional Neural Networks

CNNs have been proposed in [19] to classify 2D image data. These types of neural networks consist of several convolutional and sub-sampling layers followed by fully connected layers. An example is shown in Fig. 2. As Ravi [20] summarized, the main concepts can be expressed in three points:

- The input image is convolved using several small filters.
- The output of the previous step is sub-sampled.
- The output of the sub-sampling is considered as a new image and the convolution and the sub-sampling process is repeated until high level features are extracted.

In this way low level image features are extracted first, then by combining these low level features higher level image descriptors are obtained in the next layer. Repeating these steps, the required information is extracted from the input image to perform the classification that is the task of the last fully connected layer. The output of the last layer can be directly used to assign labels for each axial slice and to get the confidence vector of the labeling.

In this work AlexNet [8] was used for the classification of the 2D grayscale images. The structure of the NET follows the originally introduced topology.

The image dataset was split to three nearly equal subsets, such that the training, the cross-validation, and the test

sets included 66 000, 55 000, and 60 000 labeled images, respectively. The first set was used to train the classifier, the second was used to monitor the accuracy during the training, and the last one was used to evaluate the CNN model (as well as the post-processed result). In order to simulate all possible patient positions, the axial slices of the training dataset (that included images acquired in supine position only) were rotated 0, 90, 180, and 270 degrees. The training was performed within Nvidia DIGITS framework [21] using the Caffe [22] deep learning toolkit. Stochastic Gradient Descent method was applied as a solver. 30 training epochs were performed. The learning rate was set to 0.01 at the beginning and it was divided by 10 after 10 and 20 epochs. It took 5 hours to train the CNN using a simple (Nvidia Quadro K4000) GPU.

### 7 Initial Classification

In this section the initial classification results are presented to illustrate the capabilities of the CNN, to highlight its limitations and to establish the necessity of the proposed post-processing method.

Table 3 presents the confusion matrix of the applied CNN network evaluated on the test data set. The columns of the table specify the known ground truth labeling while the rows represent the labeling assigned by the trained system.

For each label the Precision value was calculated as the number of the correctly labeled slices divided by the number of all slices that were associated with that specific label by the CNN. The Recall metric is defined as the number of the correctly labeled slices for a specific label divided by the total number of slices belonging to that specific class according to the ground truth labeling. The overall accuracy (the last cell in the last row) is defined as the ratio of

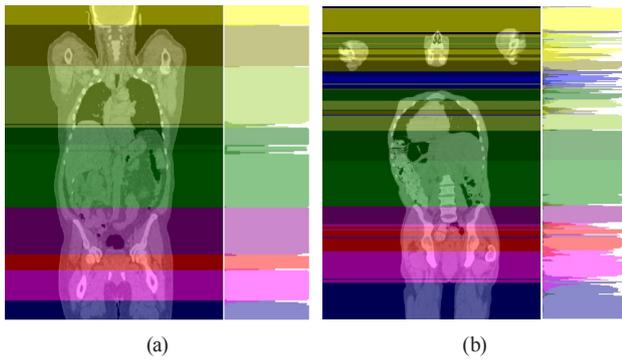
the number of all correctly labeled slices to the total number of slices in all classes.

The overall accuracy of the initial classification was 93.4 % and the precision varied between 85.1 % and 97.4 %. The values in the main diagonal represent the correctly labeled slices. The values directly below or above the main diagonal indicate some minor confusion, usually occurring at the border of the anatomical regions. For example 143 HEAD slices were labeled as CHEST. Other confusions mean serious misclassification errors. This effect indicates the independent processing of the slices, as the labeling process does not incorporate the neighborhood information at this point.

Although the majority of the slices were correctly labeled by the CNN there are some major artifacts that makes the result impractical for direct clinical application. Fig. 3 shows two examples from the test set. In the images the central coronal slice of the CT exams are shown and each row of the images represent an axial slice of the corresponding CT exam. In Fig. 3(a) the initial classification accuracy is 84.7 % and it is free from significant misclassifications while Fig. 3(b) is one of the worst initial results (accuracy: 67.8 %) and contains examples for all major errors (these errors may caused by the non-supine positioning during the acquisition). The left side of the images show the initial labeling (based on the maximal confidence values) and the confidence values are plotted on the right side. The different colors indicate different anatomical regions (yellow - BRAIN, olive - HEAD, yellow-green - CHEST, dark green - ABDOMEN-UPPER, green - ABDOMEN-LOWER, dark magenta - PELVIS-UPPER, red - PELVIS-CENTER, magenta - PELVIS-LOWER, dark blue - THIGH, middle blue - SHIN, blue - FOOT).

**Table 3** Confusion matrix of the initial labeling.

Pred\True	FOOT	SHIN	THIGH	PE-LO	PE-CR	PE-UP	AB-LO	AB-UP	CHEST	HEAD	BRAIN	Precision
FOOT	1670	94	57	0	0	7	0	3	4	9	3	90.4 %
SHIN	64	5662	78	0	0	0	0	0	2	0	7	97.4 %
THIGH	17	89	5273	318	0	2	0	0	8	15	4	92.1 %
PE-LO	0	0	133	3377	106	4	146	0	0	0	0	89.7 %
PE-CR	0	0	0	399	3782	257	0	3	4	0	0	85.1 %
PE-UP	0	0	0	51	518	15140	0	130	34	12	0	95.3 %
AB-LO	0	0	33	61	193	549	14335	801	40	11	0	89.5 %
AB-UP	21	0	15	25	0	11	239	12455	186	45	0	95.8 %
CHEST	3	0	0	4	28	11	265	436	15087	143	17	94.3 %
HEAD	0	0	0	10	0	18	39	19	67	4033	68	94.8 %
BRAIN	30	0	7	0	0	0	0	2	0	121	5968	97.4 %
Recall	92.5 %	96.9 %	94.2 %	79.6 %	81.7 %	94.6 %	95.4 %	89.9 %	97.8 %	91.9 %	98.4 %	93.4 %



**Fig. 3** Initial classification results. (a) Acceptable initial classification with minor errors, (b) Initial classification with major errors.

As the labeling of the slices are independent, the CNN can produce alternating labeling near the region borders (e.g. between PELVIS-CENTER and PELVIS-UPPER) or can completely miss the labeling as can be seen in the HEAD and CHEST regions (see Fig. 3(b)). This confusion of non-neighboring regions are major classification errors and affect 1.5 % of the slices of the whole test set.

### 8 Post-processing

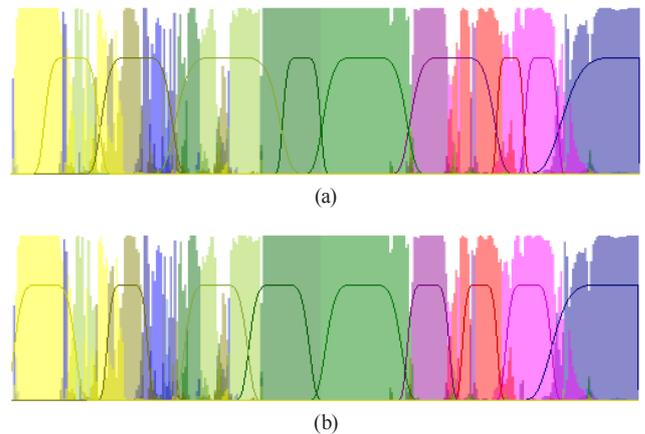
To eliminate the above mentioned artifacts of the initial classification, a post-process is introduced. The main goal of this step is to produce a continuous and reliable labeling for the whole image series. A labeling can be considered acceptable if it meets the following requirements:

- The labeling must be continuous. No alternating labeling can occur.
- The anatomical regions should appear in their correct order e.g. the CHEST region should be followed by the HEAD
- The size of the anatomical regions should be reasonable. The regions cannot be arbitrary small or large.

To obtain the desired properties of the labeling a region membership function fitting algorithm is proposed. In this method the anatomical regions are represented with generalized normal distributions (GDF) [23] as defined in Eq. (1), where  $\mu$  is the location,  $\alpha$  is the scale,  $\beta$  is the shape parameter and  $\Gamma(x)$  is the gamma function. Examples of this function can be seen in Fig. 4.

$$\text{GDF}(x) = \beta / (2\alpha\Gamma(1/\beta)) / \exp(-|x - \mu|/\alpha)^\beta \quad (1)$$

Each membership function can be customized to have region specific parameters. To estimate the position ( $\mu$ ) and range of the membership functions ( $\alpha$ ), the confidence



**Fig. 4** Membership function fitting. (a) Initialization of the membership functions, (b) Final position and range values of the membership functions.

values, provided by the CNN, can be utilized. The  $\beta$  shape parameter is set to a fixed value ( $\beta = 6$ ) in this work.

In the first step of the method, a set of regions, which are possibly involved in the image, is selected. This provides a list of candidate regions. From this list the anatomically correct sequences are assembled and the most likely one is chosen for further processing.

In the next step, a membership function set is produced to represent the selected region sequence. Based on the confidence values of the initial labeling the position ( $\beta$ ) and range ( $\alpha$ ) parameters of the functions are optimized. In this way the functions are fit on the image series and a realistic final labeling can be obtained.

#### 8.1 Selecting the Most Likely Region Combination

To estimate the visible regions in the image, the confidence weighted size of each region  $S_i$  is calculated as described in Eq. (2):

$$S_i = \sum_{k=1}^N \text{conf}(k, i) T_k, \quad (2)$$

where  $N$  is the number of the slices,  $\text{conf}(k, i)$  is the confidence value provided by the CNN describing the likelihood of that slice  $k$  is located in region  $i$ , and  $T_k$  is the physical thickness of slice  $k$ .

If  $S_i$  is greater than a predefined limit, the region is considered visible in the image. In this way the small or misclassified regions can be eliminated from the further processing.

As soon as the visible regions are available, all of the anatomically correct sequences of these regions are assembled. To select the most likely sequence the Summarized Accepted Confidences (SAC) is calculated in Eq. (3):

$$Rc_i = \left( \sum_{k=1}^N \text{conf}(k, i) \cdot k \right) / \sum_{k=1}^N \text{conf}(k, i) \quad (3)$$

$$SAC = \sum_{i=1}^M \sum_{k=1}^N \text{conf}(k, i) [Rc_{i-1} < k < Rc_{i+1}],$$

where  $M$  is the number of regions in the sequence,  $Rc_{i-1}$  and  $Rc_{i+1}$  are the confidence weighted center of the previous and next regions, respectively.  $[Rc_{i-1} < k < Rc_{i+1}]$  is 1 if  $k$  is between  $Rc_{i-1}$  and  $Rc_{i+1}$  otherwise it is 0. This means the confidence value  $\text{conf}(k, i)$  is accepted only if the position of the  $k$ -th slice is between the centers of the neighboring regions. This eliminates the most interfering outliers from the SAC calculations. This way the SAC value indicates the summarized confidence values of the slices that are positioned and labeled correctly according to the processed sequence.

Once all the SAC values are calculated, the region sequence with the highest SAC can be selected as the most likely combination of the candidate regions.

## 8.2 Fitting the Region Membership Functions

After selecting the most likely region sequence, a set of membership functions is fitted on the confidence values. This process estimates the position and range parameters of the functions to provide a continuous and reliable labeling. Regarding these calculations, the following considerations are taken:

- The membership functions should maximize the correctly covered confidence values.
- The membership functions should minimize the incorrectly covered confidence values.
- The membership functions should minimize the overlap between each other.

To perform the parameter estimation an iterative method is used. The position of the membership functions ( $\mu$ ) are initialized with the confidence weighted center of each region ( $Rc_i$ ) and their range property ( $\alpha$ ) is set to the average size of the represented region (see Table 2). Fig. 4(a) shows the initial position and range estimations for a given confidence distribution.

To meet the requirements, a cost function is defined that consists of three components. The first one, defined in Eq. (4), penalizes if a membership function ( $GDF_j$ ) does not cover slices that belong to region  $j$ .

$$cf_1 = \sum_{j=1}^M \sum_{k=1}^N (1 - GDF_j(k)) (1 + \text{conf}(k, j))^2 \quad (4)$$

The second component Eq. (5) penalizes if a membership function covers slices that do not belong to that specific region.

$$cf_2 = \sum_{j=1}^M \sum_{k=1}^N \sum_{i=1, i \neq j}^M GDF_j(k) * (1 + \text{conf}(k, i))^2 \quad (5)$$

The third component of the cost function Eq. (6) tries to minimize the overlap between the region membership functions.

$$cf_3 = \sum_{j=1}^M \sum_{k=1}^N \sum_{i=1, i \neq j}^M \min(GDF_j(k), GDF_i(k)) \quad (6)$$

The final cost function is defined as the sum of the three components.

$$cf = cf_1 + cf_2 + cf_3 \quad (7)$$

To perform the optimization and to find the desired values for the  $\alpha$  and  $\mu$  parameters of the region membership functions a gradient descent iterative optimization method is applied.

Fig. 4(b) shows the membership functions after the post-process. It can be observed that the functions are positioned according to the input confidence values and their range fit the confidences as well. The correct ordering of the regions is maintained during the iterations and their size is in the acceptable range.

## 9 Results

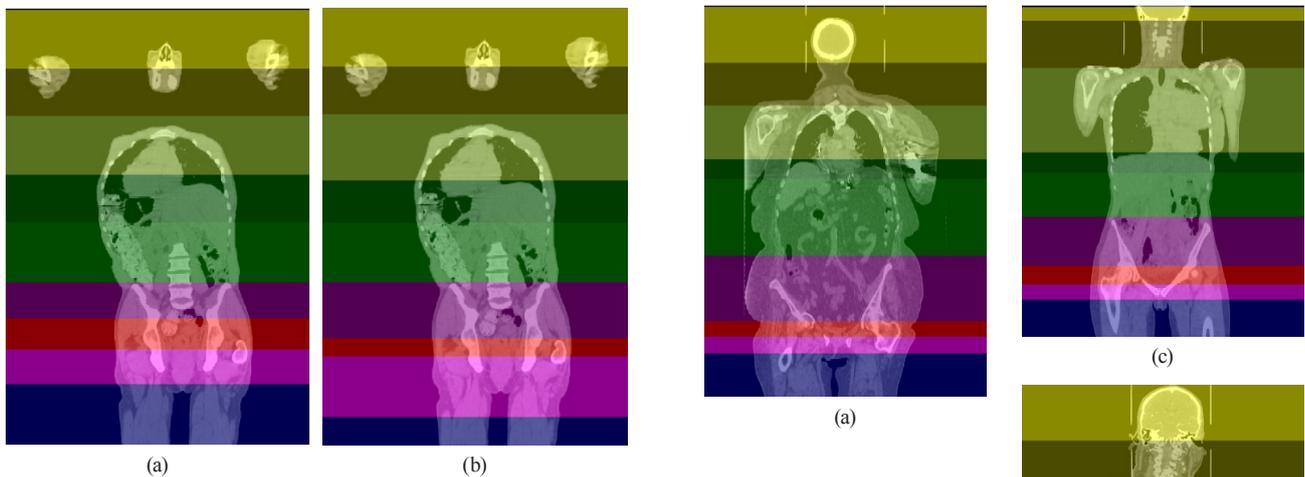
Using the post-process, the artifacts of the initial labeling can be eliminated and a continuous and reliable labeling can be obtained. The post-processed results are summarized in Table 4. As one can see the overall accuracy increased to 94.0 %. Furthermore, only the main diagonal and values below and above that are not zero, so the most interfering errors, confusion of non-neighboring regions, were completely eliminated thanks to the post-process. As the post-process guarantees the continuous labeling, the values directly below and above the main diagonal of the confusion matrix mean only small displacements of the borders between the neighboring regions.

To illustrate the effect of the post-process Fig. 5(a) is presented. Compared to Fig. 3(b) one can see, that the initial misclassification errors are eliminated and a continuous labeling is obtained without alternating labels near to the borders of the regions. Fig. 5(b) shows the ground truth labeling of the examined CT series. It can be observed that the order of the anatomical regions is correct, most of the regions have the correct size and only the PELVIS-LOWER region became smaller than the original one in the ground truth image.

In Fig. 6 other series can be seen to demonstrate the accuracy of the presented method. It can be observed that the method works well in challenging cases. Fig. 6(a)

**Table 4** Confusion matrix of the post-processed labeling.

Pred.\True	FOOT	SHIN	THIGH	PE-LO	PE-CR	PE-UP	AB-LO	AB-UP	CHEST	HEAD	BRAIN	Precision
FOOT	1769	138	0	0	0	0	0	0	0	0	0	92.8 %
SHIN	36	5616	113	0	0	0	0	0	0	0	0	97.4 %
THIGH	0	91	5419	394	0	0	0	0	0	0	0	91.8 %
PE-LO	0	0	64	3264	152	0	0	0	0	0	0	93.8 %
PE-CR	0	0	0	490	3942	279	0	0	0	0	0	83.7 %
PE-UP	0	0	0	97	533	15292	214	0	0	0	0	94.8 %
AB-LO	0	0	0	0	0	428	14525	887	0	0	0	91.7 %
AB-UP	0	0	0	0	0	0	305	12431	306	0	0	95.3 %
CHEST	0	0	0	0	0	0	0	531	15088	249	0	95.1 %
HEAD	0	0	0	0	0	0	0	0	38	4025	89	96.9 %
BRAIN	0	0	0	0	0	0	0	0	0	115	5978	98.1 %
Recall	98.0 %	96.1 %	96.8 %	76.9 %	85.2 %	95.6 %	96.6 %	89.8 %	97.8 %	91.7 %	98.5 %	94.0 %

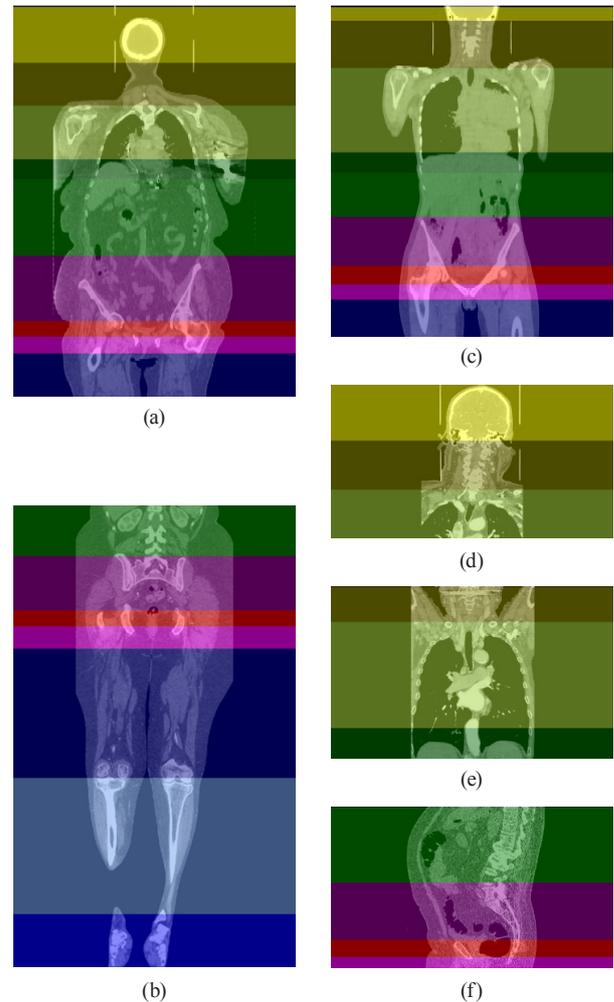


**Fig. 5** Post-processed result. (a) Corrected labeling. Accuracy: 84.5 %, (b) Ground truth labeling.

shows a patient with high level of obesity and the image has an artifact in the right arm, Fig. 6(b) is a long leg section, in Fig. 6(c) a large tumor can be seen in the left lung, Fig. 6(d) and (e) are partial body scans that are frequently acquired in the clinical practice while Fig. 6(e) shows a scan with non-supine patient position during the acquisition.

From the above mentioned examples one can see that the presented method can handle wide range of clinical cases reliably and can produce continuous labeling.

The required time to perform the calculations is composed of two factors. The initial slice classification takes a few (2-3) seconds to run using GPU, while the post-processing step takes an additional two seconds to run. In summary, the presented method takes four to five seconds to process an average CT examination.



**Fig. 6** Examples for different imaging challenges. (a) High level of obesity and imaging artifact. Accuracy: 94.6 %, (b) Partial body scan. Long leg region. Accuracy: 96.7 %, (c) Pathological distortion in the lung. Accuracy: 94.4 %, (d) Partial body scan. Head and neck Accuracy: 93.9 %, (e) Partial body scan. Chest. Accuracy: 94.7 %, (f) Non-supine position (laying on the right side). Accuracy: 85.4 %.

## 10 Conclusion

In conclusion, this paper presents a method that detects anatomy regions in CT images. The proposed approach applies deep-learning to classify 2D axial slices of the image followed by a post processing step that incorporates a-priori anatomy information to improve the accuracy. The major advantages of the method is that it recognizes all anatomy regions, it can process whole-body as well as partial scans, and it does not require the whole cross-section of the body to be involved in the image. The evaluation shows that the proposed technique provides accurate results on wide range of clinical CT cases, so it

can be a good basis of an automated image processing system. In future work the method can be extended to other 3D imaging modalities such as Magnetic Resonance.

## Acknowledgement

This work was supported by the research projects OTKA K-124124, VKSZ-14 SCOPIA, EFOP-3.6.2-16-2017-00013 and Analytic Healthcare Quality User Information Program of the National Research, Development and Innovation Fund, Hungarian Government, Grant VKSZ 12-1-2013-0012.

## References

- [1] Jiménez del Toro, O. A., Goksel, O., Menze, B., Müller, H., Langs, G., Weber, M.-A., Eggel, I., Gruenberg, K., Holzer, M., Jakab, A., Kontokotsios, G., Krenn, M., Salas Fernandez, T., Schaer, R., Taha, A. A., Winterstein, M., Hanbury, A. "VISCERAL — VISual Concept Extraction challenge in RAdioLogY: ISBI 2014 Challenge Organization", In Goksel, O. (ed.), Proceedings of the VISCERAL Challenge at ISBI, Beijing, China, 2014, pp. 6–15. [online] Available at: <http://ceur-ws.org/Vol-1194/visceralISBI14-0.pdf> [Accessed: 05 March 2017]
- [2] Golan, R., Jacob, C., Denzinger, J. "Lung Nodule Detection in CT Images using Deep Convolutional Neural Networks", In: 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 2016, pp. 243–250. <https://doi.org/10.1109/IJCNN.2016.7727205>
- [3] Bevilacqua, V. "Three-dimensional virtual colonoscopy for automatic polyps detection by artificial neural network approach: New tests on an enlarged cohort of polyps", *Neurocomputing*, 116, pp. 62–75, 2013. <https://doi.org/10.1016/j.neucom.2012.03.026>
- [4] Abdolmanafi, A., Duong, L., Dahdah, N., Cheriet, F. "Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography", *Biomedical Optics Express*, 8(2), pp. 1203–1220, 2017. <https://doi.org/10.1364/BOE.8.001203>
- [5] Pereira, S., Pinto, A., Alves, V., Silva, C. A. "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images", *IEEE Transactions on Medical Imaging*, 35(5), pp. 1240–1251, 2016. <https://doi.org/10.1109/TMI.2016.2538465>
- [6] Sharma, H., Zerbe, N., Klempert, I., Hellwich, O., Hufnagl, P. "Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology", *Computerized Medical Imaging and Graphics*, 61, pp. 2–13, 2017. <https://doi.org/10.1016/j.compmedimag.2017.06.001>
- [7] Gueld, M. O., Kohnen, M., Keysers, D., Schubert, H., Wein, B. B., Bredno, J., Lehmann, T. M. "Quality of DICOM header information for image categorization", *Proceedings of SPIE*, 4685, pp. 280–287, 2002. <https://doi.org/10.1117/12.467017>
- [8] Krizhevsky, A., Sutskever, I., Hinton, G. E. "ImageNet Classification with Deep Convolutional Neural Networks", In: NIPS 2012: Neural Information Processing Systems Conference, Lake Tahoe, Nevada, USA, pp. 1106–1114, 2012.
- [9] Szegedy, C., Toshev, A., Erhan, D. "Deep Neural Networks for Object Detection", *Nips 2013*, pp. 1–9, 2013.
- [10] Wei, Y., Xia, W., Lin, M., Huang, J., Ni, B., Dong, J., Zhao, J., Yan, S. "HCP: A Flexible CNN Framework for Multi-Label Image Classification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(9), pp. 1901–1907, 2016. <https://doi.org/10.1109/TPAMI.2015.2491929>
- [11] Haas, B., Coradi, T., Scholz, M., Kunz, P., Huber, M., Oppitz, U., André, L., Lengkeek, V., Huyskens, D., van Esch, A., Reddick, R. "Automatic segmentation of thoracic and pelvic CT images for radiotherapy planning using implicit anatomic knowledge and organ-specific segmentation strategies", *Physics in Medicine & Biology*, 53(6), pp. 1751–1771, 2008. <https://doi.org/10.1088/0031-9155/53/6/017>
- [12] Pauly, O., Glocker, B., Criminisi, A., Mateus, D., Möller, A. M., Nekolla, S., Navab, N. "Fast Multiple Organ Detection and Localization in Whole-Body MR Dixon Sequences", In: Fichtinger, G., Martel, A., Peters, T. (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011*. MICCAI 2011. Lecture Notes in Computer Science, Vol. 6893, Springer, Berlin, Heidelberg, 2011, pp. 239–247. [https://doi.org/10.1007/978-3-642-23626-6\\_30](https://doi.org/10.1007/978-3-642-23626-6_30)
- [13] Seifert, S., Barbu, A., Zhou, S. K., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D. "Hierarchical parsing and semantic navigation of full body CT data", *Proceedings of SPIE*, 7259, pp. 725–7259–8, 2009. <https://doi.org/10.1117/12.812214>
- [14] Chen, G. H., Wachinger, C., Golland, P. "Sparse Projections of Medical Images onto Manifolds", In: Gee J.C., Joshi S., Pohl K.M., Wells W.M., Zöllei L. (eds) *Information Processing in Medical Imaging*. IPMI 2013: Lecture Notes in Computer Science, Vol. 7917, Springer, Berlin, Heidelberg, Germany, 2013, pp. 292–303. [https://doi.org/10.1007/978-3-642-38868-2\\_25](https://doi.org/10.1007/978-3-642-38868-2_25)

- [15] Wachinger, C., Mateus, D., Keil, A., Navab, N. "Manifold learning for patient position detection in MRI", In: 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Rotterdam, Netherlands, 2010, pp. 1353–1356.  
<https://doi.org/10.1109/ISBI.2010.5490248>
- [16] He, B., Huang, C., Jia, F. "Fully Automatic Multi-Organ Segmentation Based on Multi-Boost Learning and Statistical Shape Model Search", In VISCEARL@ISIB 2015 VISCERAL Anatomy3 Organ Segmentation Challenge, New York, NY, USA, 2015, p. 18.
- [17] Yan, Z., Zhan, Y., Peng, Z., Liao, S., Shinagawa, Y., Metaxas, D. N., Zhou, X. S. "Bodypart recognition using multi-stage deep learning", In: Ourselin S., Alexander D., Westin CF., Cardoso M. (eds) Information Processing in Medical Imaging. IPMI 2015: Lecture Notes in Computer Science, Vol. 9123, Springer, Cham, Switzerland, pp. 449–461, 2015.  
[https://doi.org/10.1007/978-3-319-19992-4\\_35](https://doi.org/10.1007/978-3-319-19992-4_35)
- [18] Roth, H. R., Lee, C. T., Shin, H.-C., Seff, A., Kim, L., Yao, J., Lu, L., Summers, R. M. "Anatomy-specific classification of medical images using deep convolutional nets", In: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), New York, NY, USA, 2015, pp. 101–104.  
<https://doi.org/10.1109/ISBI.2015.7163826>
- [19] Lecun, Y., Bottou, L., Bengio, Y., Haffner, P. "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, 86(11), pp. 2278–2324, 1998.  
<https://doi.org/10.1109/5.726791>
- [20] Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., Yang, G.-Z. "Deep Learning for Health Informatics", *IEEE Journal of Biomedical and Health Informatics*, 21(1), pp. 4–21, 2017.  
<https://doi.org/10.1109/JBHI.2016.2636665>
- [21] Nvidia "DIGITS Framework" 2017, [online] Available at: <https://developer.nvidia.com/digits> [Accessed: 01 January 2017]
- [22] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T. "Caffe: Convolutional Architecture for Fast Feature Embedding", In: MM'14 Proceedings of the 22nd ACM international conference on Multimedia, Orlando, Florida, USA, 2014, pp. 675–678.  
<https://doi.org/10.1145/2647868.2654889>
- [23] Nadarajah, S. "A generalized normal distribution", *Journal of Applied Statistics*, 32(7), pp. 685–694, 2005.  
<https://doi.org/10.1080/02664760500079464>