BD2Decide

Big Data and models for personalized Head and Neck Cancer Decision support

TITLE	**Tumor and lymphnode segmentation fetures**		
Deliverable No.	D3.2		
EDITOR	Eleni Mantziou (ATC), Vasilis Tountopoulos (ATC), Leonidas Kallipolitis (ATC)		
Contributors	Alessio Fioravanti (UPM), Liss Hernández (UPM), Laura Lopez (UPM), Giuseppe Fico (UPM), Maria Teresa Arredondo (UPM), Florian Jung (Fraunhofer), Luca Mainardi (POLIMI), Valentina Corino (POLIMI), Eros Montin (POLIMI), Adriana Berlanga (MAASTRO), Giacomo Feliciani (MAASTRO), Ralph Leijenaar (MAASTRO), Mario Silva (UNIPR), Elena Martinelli (AOP),		
WorkPackage No.	WP3	**WorkPackage Title**	Requirements
Status[1]	FINAL	**Version No.**	1.0
Dissemination level	CO		
DOCUMENT ID	D3.2 Tumor and lymphnode segmentation features		
FILE ID	BD2Decide D3.2		

[1] Status values: TOC, DRAFT, FINAL

Distribution List

Organization	Name of recipients
AOP	T. Poli, E.M. Silini, E. Martinelli, G. Chiari, C. Caminiti, S. Rossi,
VUMC	R. H. Brakenhoff, H. Berkhof, P. van de Ven, T. Klausch
UDUS	K. Scheckenbach
INT	G.Curzio, L. Licitra, L. De Cecco, G. Calareso, G. Gatta, A. Trama, G. Calareso
UPM	G. Fico, A. Fioravanti, L. Hernández, L. Lopez, M.T. Arredondo
POLIMI	L. Mainardi, E. Montin, V. Corino
Fraunhofer	F. Jung, S. Wesarg
ATC	V. Tountopoulos, L. Kallipolitis, E. Mantziou
MAASTRO	P. Lambin, F. Hoebers, A. Berlanga, G. Feliciani, R. Leijenaar
AII	A. Algom, A. Kariv, R. Shefi
MME	S. Copelli, F. Mercalli
UNIPR	M. Silva, D. Lanfranco
European Commission	Project Officer: and all concerned E.C. appointed personnel and external experts

Revision History

Revision no.	Date of Issue	Author(s)	Brief Description of Change
0.7	15.12.2016	F.Jung	First draft
0.9	01.03.2017	F. Jung	Final version
1.0	30.04.2017	F. Jung	Final version approved by coordinator

Addressees of this document

This document is addressed to the BD2Decide Consortium and provides some insight on the algorithms developed, that will be integrated in the image analysis software (D3.3). It describes the algorithms in some detail, explains paths taken during the course of the work package and finishes with an evaluation of the newly developed approach.

This document will be delivered to the European Commission.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abbreviations and definitions

MRI	Magnetic resonance imaging
CT	Computed tomography
Segmentation	Partitioning an image in foreground (structure of interest) and background (rest)
Binary Mask	Data representation of a segmentation of a medical image. The foreground has value 1 and background 0
Voxel	Pixel equivalent in 3D image space
Midsagittal plane	The sagittal plane divides the body into a left and a right side

Table 1: Abbreviations table

Abstract

One crucial aspect of the BD2Decide project, is the analysis of the patient's medical image data and its' analysis. For this task the clinical partners will use an image analysis software, which will be developed by Fraunhofer IGD in the course of the BD2Decide project. Using the image analysis software, the radiologists will prepare segmentations of structures of interest (the tumor and suspicious lymph nodes) and extract relevant features from the image data using the software. These tasks are extremely time-consuming and exhausting. In order to facilitate this process and make it less error-prone, some automatic algorithms have and will be developed, to lower the clinicians' workload, speed-up the processing time and reduce the amount of user interaction needed.

1 ABOUT THIS DOCUMENT

1.1 Introduction and scope

This deliverable is the second deliverable of WP3 and describes the work done in WP3.2. This deliverable will give insight into the algorithms developed and used, as well as some reasoning about paths taken and decisions made, in the course of this work package.

The novel tumor detection and segmentation method for MRI, developed at Fraunhofer in the course of this work package will be introduced. Therefore, the methods used for this algorithm will be explained, as well as the workflow of the algorithm is described. Finally, the algorithm will be evaluated on clinical data. This will be appended to this deliverable, once the algorithm has been implemented in the image analysis software and has been distributed to and been evaluated by the clinical partners (D3.3 is due month 18, then the final version of the image analysis software will be released).

1.2 Structure of the deliverable

The deliverable is structured as follows:

Chapter 2 starts with a motivation on the necessity of an algorithm as it is developed in the course of this WP and gives some insight on the idea behind the algorithm.

Chapter 3 contains a short recap of the algorithm developed in WP3.1, which is used as basis for the method.

Chapter 4 describes the automatic tumor segmentation algorithm in detail.

The deliverable closes with the evaluation in Chapter 5.

2 MOTIVATION AND APPROACHING THE PROBLEM

When it comes to the analysis of medical image data, a computer algorithm still lacks a major ability of a clinical expert, the possibility to put the data into anatomical context. The program is not able to distinguish for a specific voxel of the image data to which corresponding structure it belongs. Depending on the structure to analysis this can be a very easy task for a clinical expert or very challenging as well. Many elements influence the difficulty of this task.

- The image modality
- The contrast between the structure of interest and surrounding structures
- The homogeneity of the structure
- The image resolution

In order for a computer algorithm to be able to further analyze (axis, volume, radiomics features) a specific structure, a segmentation of the structure is needed. A segmentation is a mask/overlay of the original image data that defines which of the image voxels belongs to the structure of interest and which don't. As mentioned before, even for a clinical expert this can be very challenging and it is always an extremely time consuming task.

Experiences from the OraMod project (FP7-2013-486715) have shown, that the average time needed for a tumor segmentation lies between 15 and 30 minutes. As a result, the segmentation of the tumor alone needs more time than all other tasks combined, which are performed by the clinical expert in the course of a single patient's analysis.

Therefore, in collaboration with the clinical partners it was decided to develop an algorithm for the automatic segmentation of the tumor on the medical image data. Due to the fact that, the segmentation of a tumor takes significantly more time then the segmentation of a lymph node and Fraunhofer already possess methods to achieve a semi-automatic segmentation of lymph nodes, it was decided to solely focus the work on the tumor segmentation algorithm in this workpackage.

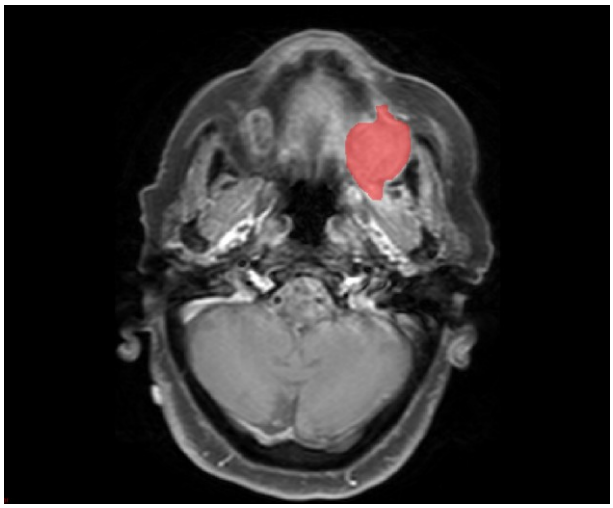

Figure 2-1: Example of tumor segmentation

Common segmentation algorithms usually are supervised by a clinical expert. This has some disadvantages. In order to start the segmentation the algorithms first needs some user input, as an initialization. Then the algorithm determines the segmentation and finally the clinician has to verify it. So the clinician has one task to complete and has to wait for the final result of the algorithm. Therefore, it would be major improvement if the algorithm would work without any manual user initialization, but instead fully automatically.

This is an extremely complicated task and to the best of our knowledge, no approach exists, which does the segmentation of tumors in the head and neck area in a fully automatic manner. This has different reasons, but mainly the issues mentioned earlier in this section, which make a segmentation difficult. One of the most crucial problems for such an approach, is the difficulty to even estimate, where within the image the tumor is located.

In order to tackle this problem, we want to exploit the human anatomy. More specific, we want to benefit from the fact, that the human head and neck area is mostly symmetrical. Since most tumors originate on one side of the body, this should result in inconsistencies in the image data of the head and neck area. A requirement would be, that the size of the tumor is large enough to result in detectable inconsistencies within the image data. This should be true for most tumors with a T-Staging of T3 or T4, which will represent the majority of all tumors analyzed within the project. Not to mention, that the segmentation of small tumors only takes the fraction of the time the larger ones need anyway.

So in order to accomplish the task, two things are needed:

- The calculation of the symmetry axis
- The detection and segmentation of the tumor using the symmetry axis

This first task has already been accomplished in deliverable 3.1 and this deliverable will focus on the second task.

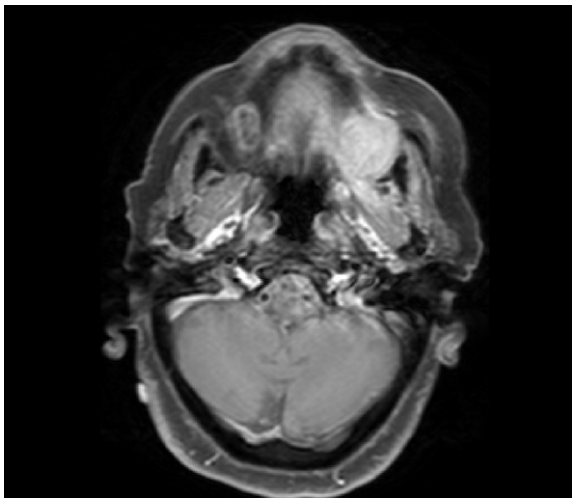

Figure 2-2: Example slice of a patient's data set with a tumor

3 METHODS

This section focuses on methods used for the realization of the automatic tumor detection and segmentation approach.

3.1 Optimization based sagittal midline determination method

This section gives a very short recap of the algorithm developed in WP3.1. To give a better understanding on want our new approach is based on.
The determination of the patient's midplane is designed as an optimization task and has the following steps:

1. Load medical image data
2. Initialize mid-sagittal plane in medical image data
3. Start optimization process:
    a. Calculate metric value for current configuration (Consider stop criterion)
    b. Optimizer proposes new parameter set for midplane
    c. The plane is transformed using new parameters
    d. Continue with step 3a) until stop criterion is met.
4. Save sagittal-midline for preceding steps

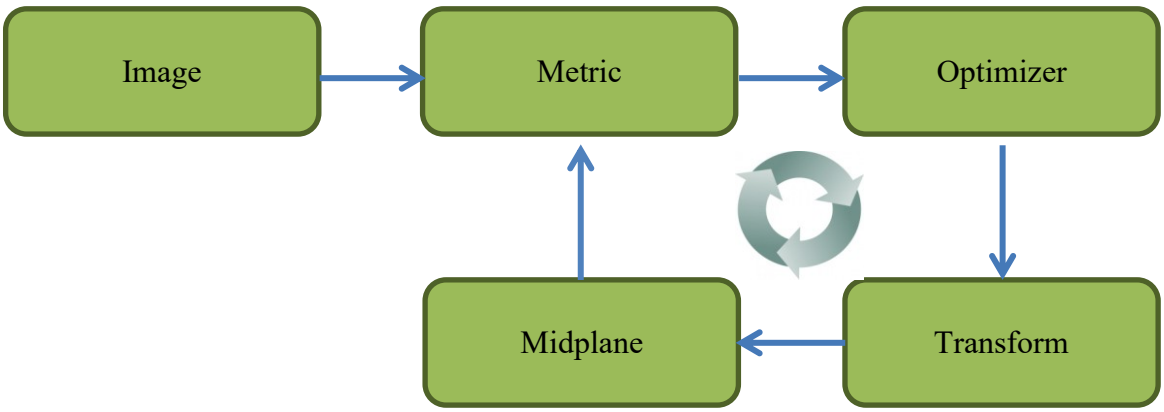

Figure 3-1: Midplane optimization process flowchart

4 AUTOMATIC TUMOR SEGMENTATION ALGORITHM

Within this chapter the automatic tumor segmentation algorithm is described in detail. The complete algorithm consists of several steps, which are visualized in consecutive order in Figure 4-1. The single steps will be explained in more detail in the different sections of this chapter.

The input data of the algorithm should be T1 weighted MRI data. The algorithm should be able to perform on other modalities as well, but some steps of the algorithm might need some minor modifications.

The output of the algorithm is a binary mask of the original data, which represents the found tumor structure.

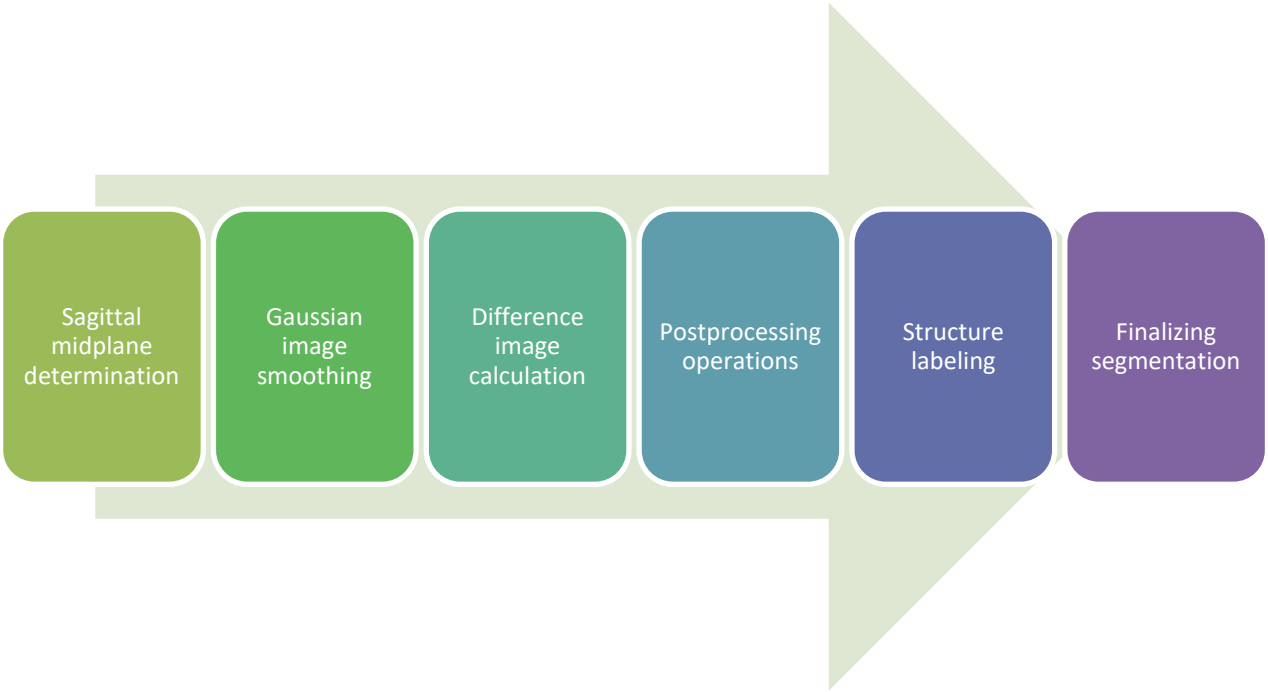

Figure 4-1: Segmentation algorithm workflow

4.1 Midsaggital plane determination

As a first step the midsaggital plane has to be found, which will be the basis for all further calculations. This algorithm has been described in detail in deliverable 3.1, which should be used as a reference for further information.

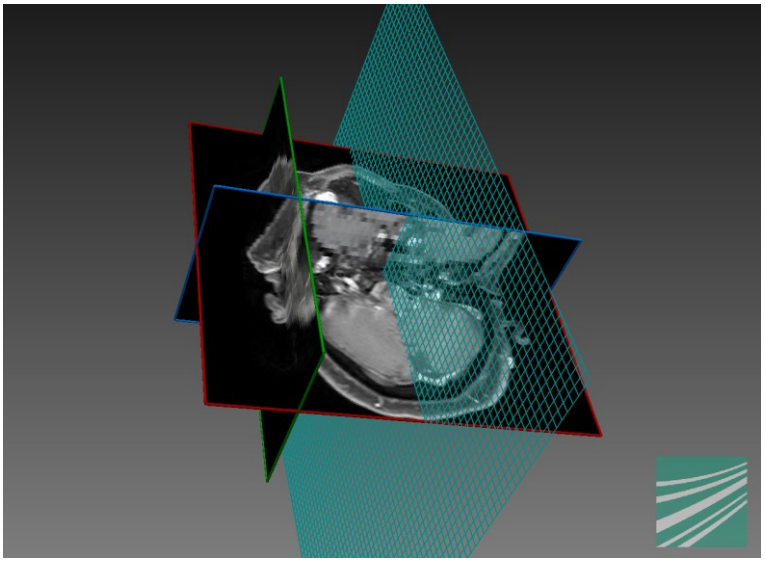

Figure 4-2: Example of an extracted midsaggital plane

4.2 Gaussian image smoothing

The second step is the usage of a Gaussian smoothing[2]. Due to the fact, that the human body, is not completely symmetrical and it can't be expected, that the midplane is always positioned perfectly within the image data, such a smoothing of the data can be extremely beneficial. Moreover, this leads to a more robust algorithm towards image artifacts. Since the main goal is to locate complete regions of the image, where it differentiates from the other side and not single voxels which differ from their counterparts on the other side of the image plane, we used 2 sigma as standard deviation for the Gaussian smoothing. This delivers a good tradeoff between smoothing and loss of information. Figure 4-3 shows the original image data as well as the smoothed image.

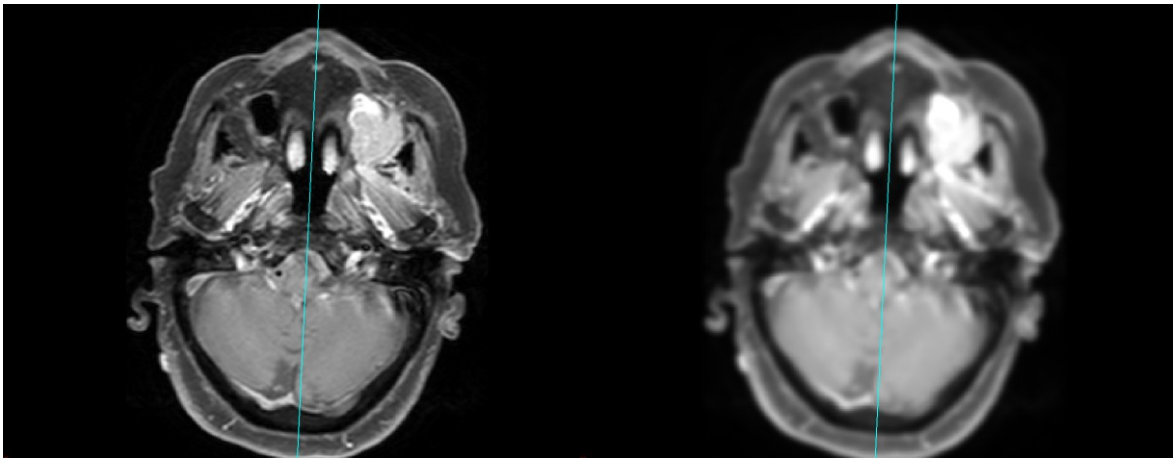

Figure 4-3: Gaussian smoothing of the image data

[2] https://en.wikipedia.org/wiki/Gaussian_blur

4.3 Difference image calculation

On the smoothed image a difference image is calculated. This is the most crucial step of the algorithm. It classifies the voxels which diverge significantly from the ones on the other side of the image plane. So this step does the actual tumor detection within the image data.

The algorithm iterates over every pixel of the image. For this it calculates the voxel position on the other side of the image plane using the following formula.

$$\vec{x}_m = \vec{x} - 2 * \vec{n} * (\vec{n} \cdot \vec{x} - \vec{n} \cdot \vec{r})$$

$\vec{x}$ is the position of the initial point. $\vec{n}$ is the normal vector of the plane. The calculation within the bracket calculates the distance of the point to the midplane, which then is multiplied with the normal vector twice, to receive the mirrored coordinate.

As a next step the Intensity difference of the 2 coordinates is calculated, but only positive values are kept, because the tumor always has higher intensity than it's corresponding tissue on the other body side.

$$d = \min(0, I(x) - I(x_m))$$

Once the calculations have been run for every voxel of the input image, we obtain a difference image, as shown in Figure 4-4. This image will be used as input data for the next processing steps, to extract the tumor location and create a segmentation of before mentioned.

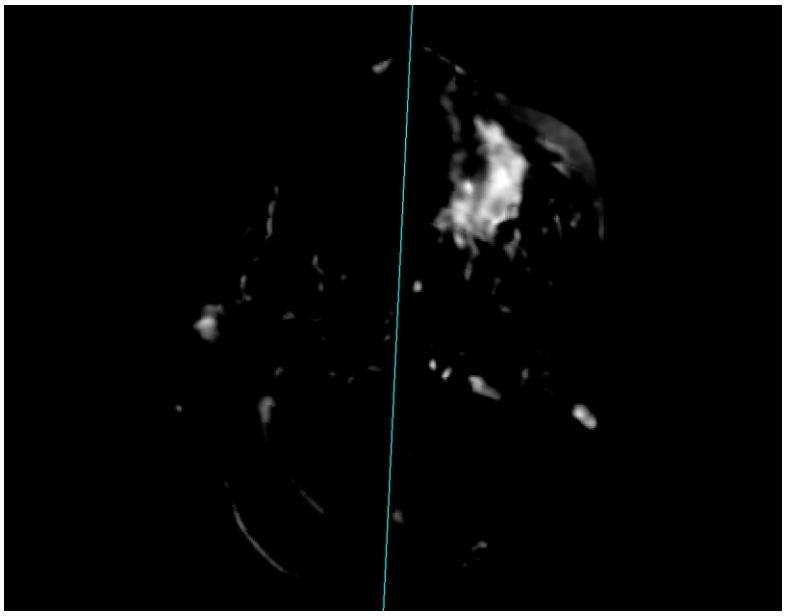

Figure 4-4: Difference image

4.4 Postprocessing operations

To further distinguish between the actual tumor and image artefacts or normal asymmetries, which have been detected up to this point, some postprocessing steps are run. After the inspection of

several tumors in different patient data sets and the analysis of their difference value histograms, it became clear, that the tumor's difference values are among the higher ones. This led us to the idea, to further filter the data by using a thresholding algorithm. To achieve this, we extract the statistical median of the difference image, where only non-zero values are considered. Next the median is used as a threshold for the difference, leading to a difference image, where all difference values that are lower than the median, are set to zero (see Figure 4-5).

The difference image is then converted to a binary image, see Figure 4-6.

Finally, a morphological opening [3]is executed on the difference image to remove small artefacts/structures from the image data (Figure 4-7).

Figure 4-5: Difference image before and after the thresholding

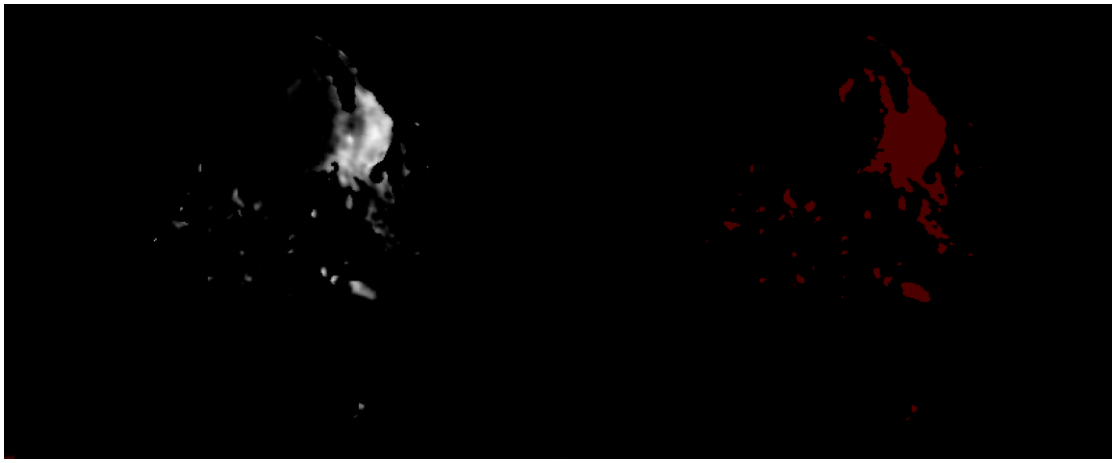

Figure 4-6: Conversion of the difference image to a binary image

[3] https://en.wikipedia.org/wiki/Opening_(morphology)

Figure 4-7: Binary image before and after the morphologic opening

4.5 Structure labeling

To do the actual tumor detection we use a filter [4]to convert the binary image to a label map. Given the binary map as input the filter calculates an output labelmap which consists of n labels. Internally it uses a connected components filter. So for every connected component a separate label in the labelmap is generated. These labels are then used to extract the final segmentation of the tumor.

4.6 Finalizing segmentation

The labels are sorted by size and the three biggest objects are saved and will be proposed as potential tumors to the clinician. The has the big advantage over only picking one segmentation and throwing the others away, that in case another big none tumor structure has been detected, the clinician still can pick the correct 2nd or 3rd largest structure (Although, until now, in all data sets we have used for testing purposes, the largest structure always was the primary tumor).

[4] http://www.insight-journal.org/browse/publication/176

5 EVALUATION

The evaluation of the newly developed segmentation algorithm will be covered in this chapter. Since the algorithm will be part of the image analysis software, which is deliverable 3.3, these evaluation results only then can be integrated, when the clinical centers have received the final version of the software. Once these results are available, this document will be updated with this information.

Until then we have run some local evaluation outside of clinical practice to acquire some first impressions of the algorithms capability to detect and segment the tumor.

The first results look very promising. On 10 tested data sets, in all cases the tumor was located correctly within the image data.

Preceding 5 example images of the segmentations from 5 different patients are shown. 4 of the 5 segmentations look already very good. Only Figure 5-2 shows an example where the segmentation result can be considered as problematic. In this case the tumor is connected to a structure with exactly the same intensity. This leads the tumor segmentation to leak into the other structure.

Summarizing it can be said, that the results are already very promising. But the actual evaluation in clinical practice has to be carried out first to give a final conclusion, how much time and work can be saved by using the newly developed algorithm. An evaluation report (ref. example Table 2) will be collected to assess the performance of the tool.

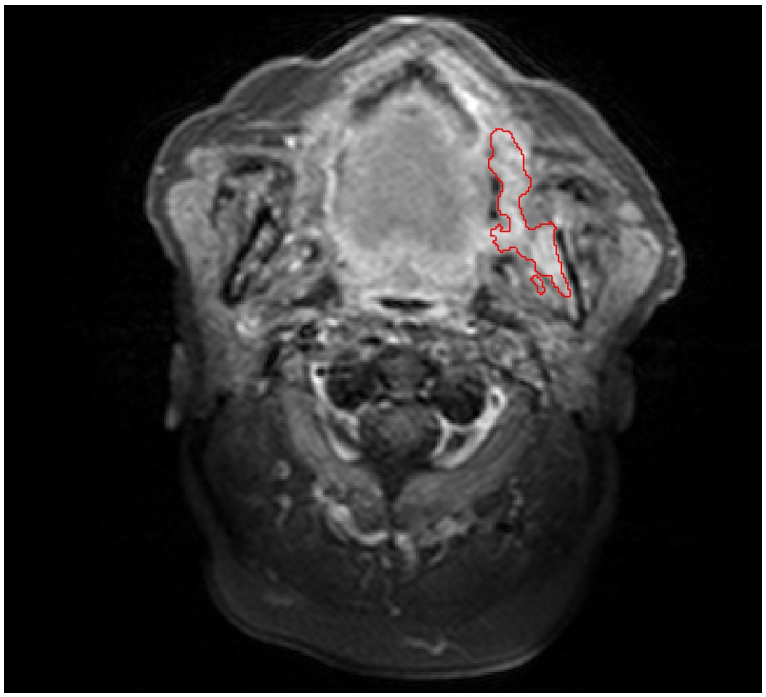

Figure 5-1: Segmentation example of patient Parma 126

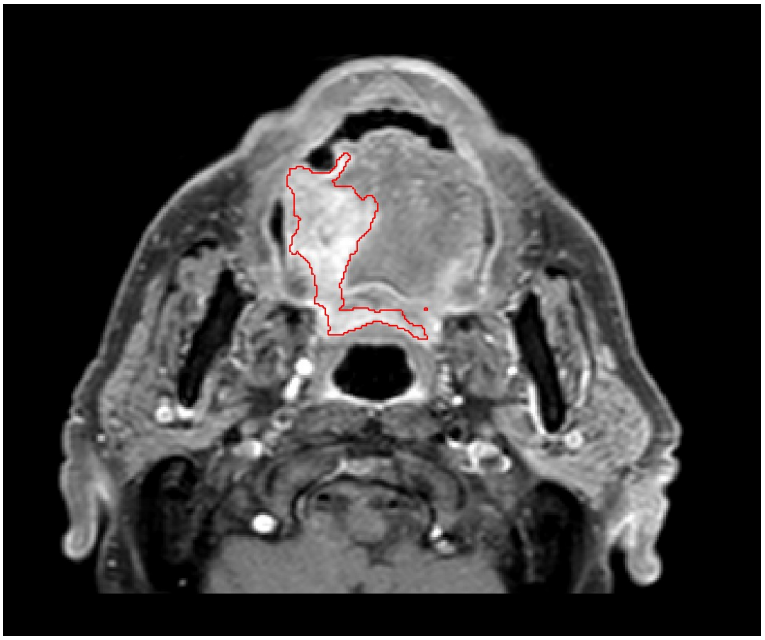

Figure 5-2: Segmentation example of patient Parma 127

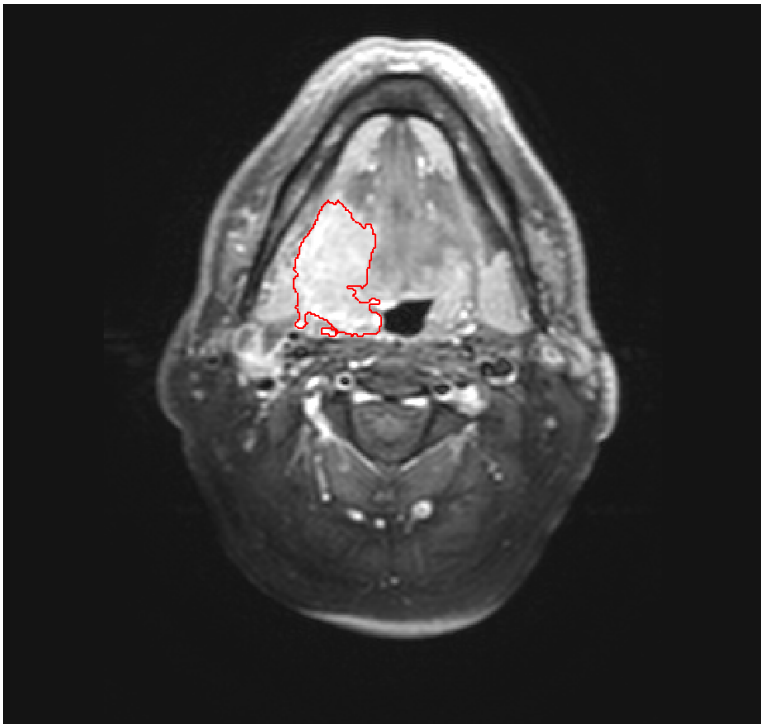

Figure 5-3: Segmentation example of patient Parma 128

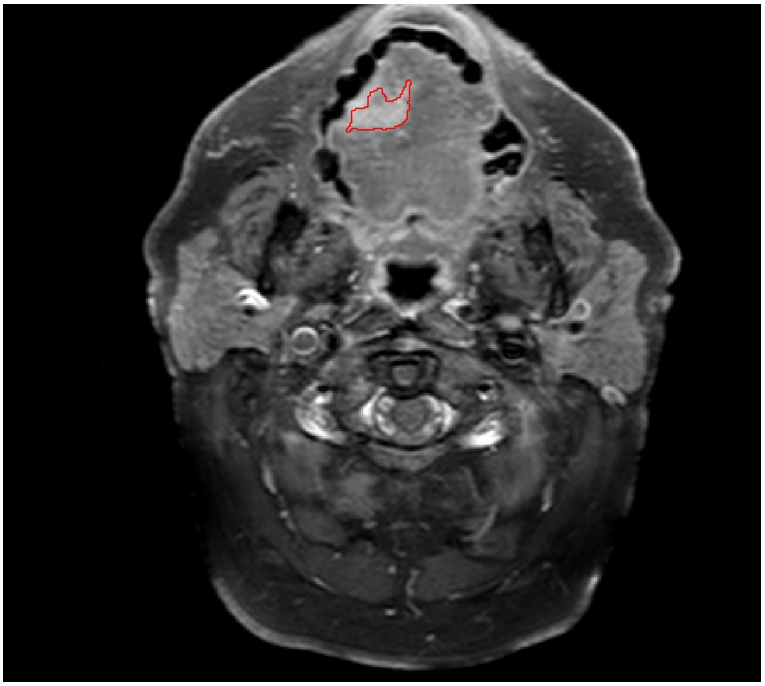

Figure 5-4: Segmentation example of patient Parma 136

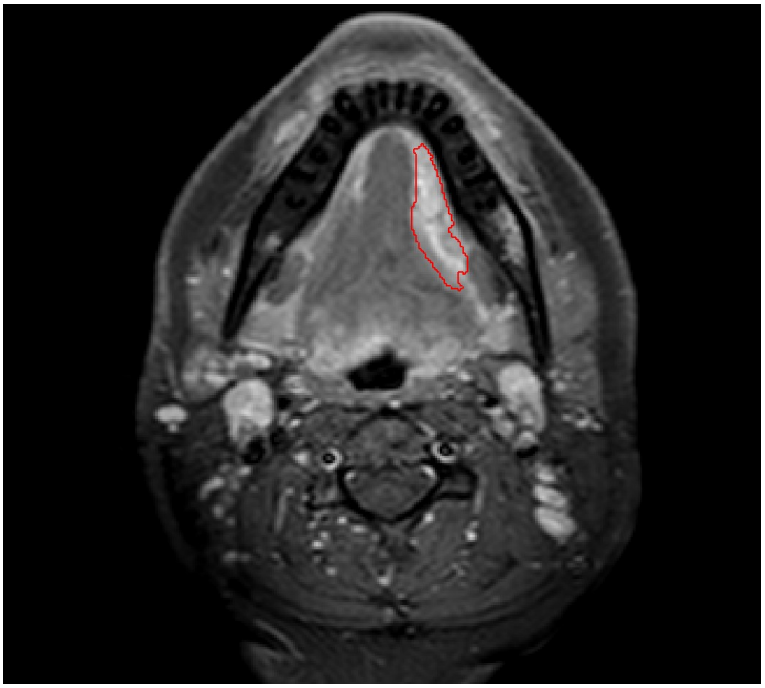

Figure 5-5: Segmentation example of patient Parma 143

Patient	Segmentation quality	Time saving
XXX	XXX	XXX

Table 2: Table of the clinical evaluation