BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

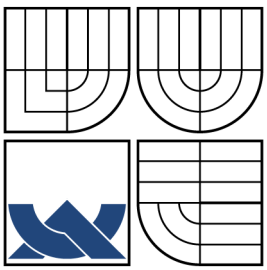
BRAIN TUMOR DETECTION AND SEGMENTATION IN MULTISEQUENCE MRI

DOCTORAL THESIS
DIZERTAČNÍ PRÁCE

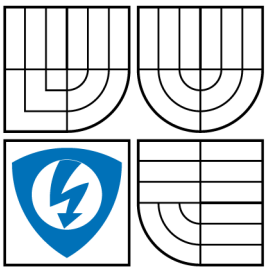
AUTHOR
AUTOR PRÁCE

Ing. PAVEL DVOŘÁK

Brno 2015



BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ



FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

BRAIN TUMOR DETECTION AND SEGMENTATION IN MULTISEQUENCE MRI

DETEKCE A SEGMENTACE MOZKOVÉHO NÁDORU V MULTISEKVENČNÍM MRI

DOCTORAL THESIS
DIZERTAČNÍ PRÁCE

AUTHOR
AUTOR PRÁCE
Ing. PAVEL DVOŘÁK

SUPERVISOR
VEDOUCÍ PRÁCE
prof. Ing. ZDENĚK SMÉKAL, CSc.

BRNO 2015

ABSTRACT

This work deals with the brain tumor detection and segmentation in multisequence MR images with particular focus on high- and low-grade gliomas. Three methods are propose for this purpose. The first method deals with the presence detection of brain tumor structures in axial and coronal slices. This method is based on multi-resolution symmetry analysis and it was tested for T1, T2, T1C and FLAIR images. The second method deals with extraction of the whole brain tumor region, including tumor core and edema, in FLAIR and T2 images and is suitable to extract the whole brain tumor region from both 2D and 3D. It also uses the symmetry analysis approach which is followed by automatic determination of the intensity threshold from the most asymmetric parts.

The third method is based on local structure prediction and it is able to segment the whole tumor region as well as tumor core and active tumor. This method takes the advantage of a fact that most medical images feature a high similarity in intensities of nearby pixels and a strong correlation of intensity profiles across different image modalities. One way of dealing with – and even exploiting – this correlation is the use of local image patches. In the same way, there is a high correlation between nearby labels in image annotation, a feature that has been used in the "local structure prediction" of local label patches. Convolutional neural network is chosen as a learning algorithm, as it is known to be suited for dealing with correlation between features.

All three methods were evaluated on a public data set of 254 multisequence MR volumes being able to reach comparable results to state-of-the-art methods in much shorter computing time (order of seconds running on CPU) providing means, for example, to do online updates when aiming at an interactive segmentation.

KEYWORDS

Brain tumor, image analysis, image segmentation, medical imaging, MRI, multisequence imaging, local structure, structure prediction

ABSTRAKT

Tato práce se zabývá detekcí a segmentací mozkového nádoru v multisekvenčních MR obrazech se zaměřením na gliomy vysokého a nízkého stupně malignity. Jsou zde pro tento účel navrženy tři metody. První metoda se zabývá detekcí prezence částí mozkového nádoru v axiálních a koronárních řezech. Jedná se o algoritmus založený na analýze symetrie při různých rozlišeních obrazu, který byl otestován na T1, T2, T1C a FLAIR obrazech. Druhá metoda se zabývá extrakcí oblasti celého mozkového nádoru, zahrnující oblast jádra tumoru a edému, ve FLAIR a T2 obrazech. Metoda je schopna extrahovat mozkový nádor z 2D i 3D obrazů. Je zde opět využita analýza symetrie, která je následována automatickým stanovením intenzitního prahu z nejvíce asymetrických částí. Třetí metoda je založena na predikci lokální struktury a je schopna segmentovat celou oblast nádoru, jeho jádro i jeho aktivní část. Metoda využívá faktu, že většina lékařských obrazů vykazuje vysokou podobnost intenzit sousedních pixelů a silnou korelaci mezi intenzitami v různých obrazových modalitách. Jedním ze způsobů, jak s touto korelací pracovat a používat ji, je využití lokálních obrazových polí. Podobná korelace existuje také mezi sousedními pixely v anotaci obrazu. Tento příznak byl využit v predikci lokální struktury při lokální anotaci polí. Jako klasifikační algoritmus je v této metodě použita konvoluční neuronová síť vzhledem k její známe schopnosti zacházet s korelací mezi příznaky.

Všechny tři metody byly otestovány na veřejné databázi 254 multisekvenčních MR obrazech a byla dosáhnuta přesnost srovnatelná s nejmodernějšími metodami v mnohem kratším výpočetním čase (v řádu sekund při použitý CPU), což poskytuje možnost manuálních úprav při interaktivní segmetaci.

KLÍČOVÁ SLOVA

Mozkový nádor, analýza obrazů, segmentace obrazů, lékařské zobrazování, MRI, multisekvenční zobrazování, lokální struktura, predikce struktury

DVOŘÁK, Pavel *Brain Tumor Detection and Segmentation in Multisequence MRI*: doctoral thesis. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Telecommunications, 2015. 113 p. Supervised by prof. Ing. Zdeněk Smékal, CSc.

DECLARATION

I declare that I have written my doctoral thesis on the theme of "Brain Tumor Detection and Segmentation in Multisequence MRI" independently, under the guidance of the doctoral thesis supervisor and using the technical literature and other sources of information which are all quoted in the thesis and detailed in the list of literature at the end of the thesis.

As the author of the doctoral thesis I furthermore declare that, as regards the creation of this doctoral thesis, I have not infringed any copyright. In particular, I have not unlawfully encroached on anyone's personal and/or ownership rights and I am fully aware of the consequences in the case of breaking Regulation § 11 and the following of the Copyright Act No 121/2000 Sb., and of the rights related to intellectual property right and changes in some Acts (Intellectual Property Act) and formulated in later regulations, inclusive of the possible consequences resulting from the provisions of Criminal Act No 40/2009 Sb., Section 2, Head VI, Part 4.

Brno ....

(author's signature)

ACKNOWLEDGEMENT

I would like to thank to the thesis supervisor prof. Ing. Zdeněk Smékal, CSc. for the guidance and professional support during the work on this thesis. I would also like to thank to prof. Ing. Karer Bartušek, CSc., Institute of Scientific Instruments of the CAS; Prof. Dr. Bjoern H. Menze, Technical University of Munich; and O.Univ.Prof. Dr. Walter G. Kropatsch, Vienna University of Technology; for their consultations and inspiring ideas.

I would like to express gratitude to my family, to my dearest girlfriend, to my colleagues and friends for their patience, understanding and their infinite support.

Brno ....

(author's signature)

Faculty of Electrical Engineering
and Communication
Brno University of Technology
Purkynova 118, CZ-61200 Brno
Czech Republic

http://www.six.feec.vutbr.cz

ACKNOWLEDGEMENT

Research described in this doctoral thesis has been implemented in the laboratories supported byt the SIX project; reg. no. CZ.1.05/2.1.00/03.0072, operational program Výzkum a vývoj pro inovace.

Brno . . . . . .

. . . . . .

(author's signature)

EVROPSKÁ UNIE
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
INVESTICE DO VAŠÍ BUDOUCNOSTI

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Based on statistics of the Central Brain Tumor Registry of the United States (CBTRUS), brain tumor is one of the leading causes of cancer-related deaths in the United States. It is the second leading cause of cancer-related deaths in children under age 20 as well as in males ages 20-39, leukemia is the first for both. Every year, an estimated number of new brain tumor diagnosis in the United States is about 70,000 in adults including malignant (24,000) and non-malignant (45,000), with an estimated number of deaths 14,000 attributed to malignant brain tumors, and 4,300 new diagnosis in children, where more than half of these are in children younger than 15. The prevalence rate for all primary brain and CNS tumors is 221.8 (61.9 for malignant and 177.3 for non-malignant) per 100,000 people. The average annual mortality rate in the US was 69,789 attributed to primary malignant brain and CNS tumors between 2007 and 2011. The five-year survival after malignant brain tumor diagnosis decreases with age with e.g. 58.5% for adults between 20 and 44 and 31.1% for adults between 45 and 54. These statistics include only primary brain tumors, those that begin and tent to stay in the brain.

The most common primary brain tumor is meningioma with 34%, however glioma, a broad term including tumors arising from the gluey or supportive tissue of the brain (30% of all brain tumors), represents 80% of malignant tumors making it the most common primary brain tumor causing death. The most common and aggressive glioma is glioblastoma multiform (GBM) representing 54% of all gliomas. This type of tumor is accompanied by rapid infiltrative growth and very poor prognosis with one year average survival time after diagnosis [116]. The survival time is affected by extensive treatment such as chemo- and radiotherapy and surgical resection. This work is particularly focused on the automatic processing of volumes with the most common malignant tumor - glioma - in low and high grades.

In common clinical routines, the evaluation of acquired images is currently performed manually based on quantitative criteria or measures such as the largest visible diameter in axial slice [32]. Therefore, highly accurate methods being able to automatically analyze scans of brain tumor would have an enormous potential for diagnosis and therapy planning. However, it was shown by Menze et al. [80] that even manual annotation performed by expert raters showed significant variations in areas where intensity gradients between tumorous structures and surrounding tissues are smooth or obscured by bias field artifacts or the partial volume effect. Moreover, brain tumor lesions are only defined by relative intensity changes to healthy tissues, and their shape, size and location are individual for each patient, which makes the use of common pattern recognition algorithms impossible.

Due to the increasing number of patients, the number of acquired data increase,

too. Therefore, there is an increasing necessity of algorithms that are able to process the data automatically, which is the main motivation for this work.

MR sequences

Magnetic Resonance Imaging (MRI) is commonly used for brain tumor analysis and exploration. A variety of MR sequences exists, where each of them is suitable for different imaging purpose. Nowadays, it is a common practice in automatic analysis to use a combination of several MR sequences to achieve more valuable and accurate results. In this work, three different MR sequences, namely T1-weighted image, T2-weighted image, and FLAIR image, are used and they will be characterized shortly as described in [4].

T1-weighted image. In MRI, T1 refers to the time that protons within a tissue need to return to the initial magnetization state, which is given by the static magnetic field. Simple T1-weighted images (shortly T1 images) provide better anatomical details than T2-weighted images but they usually do not bring interesting information when brain tumor is investigated. However, they are used in a combination with contrast agent fluid, which is injected into patient's vascular system. The contrast agent highlights the blood flow in T1-weighted images. This cause the tumor active part, as well as vessels, to appear hyper-intense and easily distinguishable from surrounding tissues. The presence of the active tumor is often used during the tumor malignancy investigation. Such images are called "contrast enhanced T1-weighted images" and in this work the abbreviation T1C will be used.

T2-weighted image. T2 refers to the time that protons perturbed into coherent oscillation by radiofrequency pulse require to lose this coherence. T2-weighted images (shortly T2 images) are, compared to T1 images, more sensitive to the content of water and, therefore, to the pathology, which, as well as cerebrospinal fluid (CSF), appears hyper-intense here.

FLAIR image. Fluid-attenuated inversion recovery (FLAIR) is a sequence that is able to suppress fluids and it is used to suppress CSF in brain imaging. This effect enables to distinguish lesions, which remains hyper-intense as in T2 images, from CSF which becomes hypo-intense here. For that reason, it is commonly used in the brain tumor imaging.

A comparison between T1, T1C, T2 and FLAIR images with tumor present is depicted in Fig. 1.1. Note the hyper-intense active tumor in T1C image, hyper

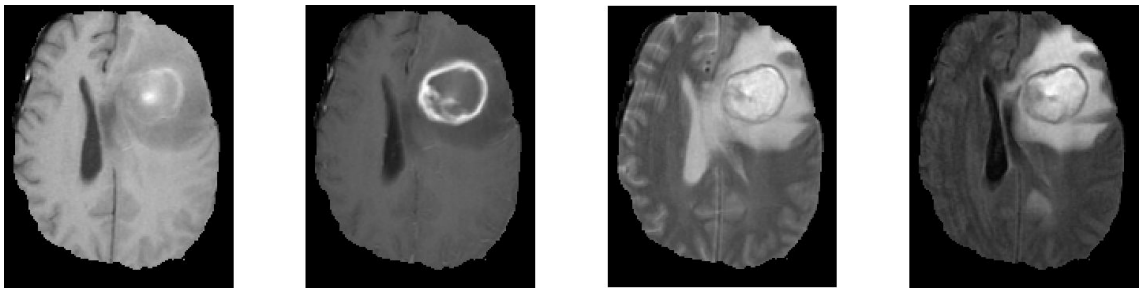


Fig. 1.1: From left to right: T1, T1C, T2 and FLAIR images of the same slice.

intense tumor and edema in T2 and FLAIR images, and hypo-intense CSF in FLAIR image.

Structure of the thesis

This dissertation is organized as follows. Chapter 2 reviews the existing related state-of-the-art works based on different approaches and describes their advantages and disadvantages. In Chapter 3, the state-of-the-art methods of local structure prediction are described. Chapter 4 describes the goals of the work and the expected contribution to the scientific world. In Chapter 5, three proposed methodologies for brain tumor detection and segmentation are presented. This chapter also includes a brief description of pre-processing of MR images. The results of the proposed frameworks as well as the experimental setup, discussion and comparison with state-of-the-art works are described in Chapter 6. Finally, the conclusion is given in Chapter 7, where the possible future work is described too.

2 RELATED WORK - BRAIN TUMOR SEGMENTATION

The aim of this chapter is to review algorithms developed for segmentation of brain tumors and surrounding tissues. The automated brain tumor segmentation is still a challenging task. One of the reason is tumor's unpredictable properties such as size, shape and location unless the tumor development in time is investigated and images from previous scanning are available. Considering only the independent scanning, all of the mentioned properties are unknown. Thus common pattern recognition techniques relying on such properties and widely used for object detection and extraction in both medical and real world images cannot be employed. However, other knowledge such as structure of the healthy human brain or tumor manifestation in particular MR sequences can be used. This is, on the other hand, the advantage compared to image object detection, e.g. human or car, where the color and the background scene vary. There have been an increasing interest in developing such algorithms and, particularly, the automatic brain tumor segmentation task has recently attracted many computer vision research teams.

Due to variety of brain tumor types and their manifestation in MR images, most state-of-the-art methods focus on most common tumor types, i.e. glioblastoma, or they require specific training database to deal with a specific tumor type. Only few researchers, e.g. Islam et al. [51], tried to train a developed algorithm on one tumor type and test it on another. However, the results were not satisfying.

Since different brain tumor segmentation methods rely on different image information, the following division will be used in this chapter: threshold-based methods, which rely on the intensity difference between a brain tumor and surrounding tissues, region-based methods, which search for connected regions of voxels with similar properties, contour-based methods searching for edges between a brain tumor and surrounding tissues, classification or clustering methods, which make use of voxel-wise intensity and texture features, and atlas-based methods using the prior knowledge about the healthy brain. However, the division is not always clear because proposed methods often have the same nature and combine more of these approaches. In recent years, most state-of-the-art techniques have been based on classification and atlas guidance.

Another possibility is to divide brain tumor segmentation methods according their requirements of a human interaction. Two groups can be defined: semi-automatic and fully-automatic. The former requires an expert interaction such as seed point selection or approximate initial boundary delineation. The latter is a group of fully automatic works without any human action. However, the former

division, i.e. according to the image information used by particular methods, will be used throughout this chapter.

2.1 Threshold-based methods

Since image intensity is the simplest property that is shared by voxels of an image region, the natural way to perform the segmentation is to set a threshold that separates a given object from other regions with voxels of different intensity. This may be reached either by single- or multi-thresholding techniques. The former may be used in case that the object intensities are clearly distinguishable and either higher or lower than background intensities, e.g. enhancing part of brain tumor core in T1C. The latter is used if intensities of a given object lie between intensities of other image objects, e.g. gray matter in T2 image. Similar effect can be reached by a cascade of single thresholds.

The thresholding assigns to every voxel a discrete value from the range $[0, t]$ where t denotes the number of thresholds. The main problem of threshold-based methods is that no relationships between voxels are considered. Therefore many voxels of other objects can be extracted, mainly in case of low signal-to-noise ratio. Image inhomogeneity across the scene making part of an object brighter or darker causes failing of such techniques too.

Such technique was used by Gibbs et al. [44] as an unsupervised segmentation technique of enhancing tumor part in T1C images, where it was applied to manually selected regions of interest as the first step of the segmentation process followed by the region growing algorithm described in the next section. However, the method does not take into account the presence of healthy hyper-intense regions in such images. Stadlbauer et al. [115] used the average intensity plus three times the standard deviation to delineate the pathological tissue from T2 images.

Another variety of image thresholding technique is based on local image properties, where the threshold value is determined adaptively according to the intensity distribution in particular image region. This is useful in cases where it cannot be estimated from the whole image or it has to be adapted for different parts of the image separately, e.g. due to the image inhomogeneity. Such technique was used by Shantni et al. [109] for brain tissue segmentation.

Generally, threshold-based techniques behave well if contrast between the object and background is high, which is not the case of brain MR images. Moreover, the threshold estimation is usually hard and requires a user interaction. However, they can be used as a first step of the segmentation process.

2.2 Region-based methods

Region-based methods are such algorithms that search for connected regions of voxels with homogeneous properties based on a predefined similarity criterion. The most famous algorithms from this group are region growing and watershed, basic image segmentation techniques, and as it will be shown in the next two subsections, they have been commonly used in the brain tumor segmentation task.

2.2.1 Region growing

Region growing belongs to the group of semi-automatic segmentation techniques since it requires a manual selection of at least one seed point, especially in the brain tumor segmentation task. However, for RGB images an automatic seeded region growing algorithm was developed by Shih and Cheng [112].

The algorithm starts in the neighborhood of the seed point and continues to more distinct areas as long as the similarity criterion is satisfied. The segmentation is finished when there are no pixels to be examined. The result of the segmentation is one connected region. Several papers ([106, 107, 124]) showed that this algorithm can be effective and less computationally intensive for brain tumor segmentation than other methods.

Weglinski and Fabijanska [127] used this algorithm for brain tumor segmentation using manual seed point selection and difference between the seed and an examined voxel as a criterion. The region growing algorithm was also used for brain tumor segmentation in MR images by Kaus et al. [54]. Their method was based on an iterative statistical classification method, which divided the input image voxels into tissue classes according to their intensity.

However, the disadvantage of region growing is its inaccuracy in case of partial volume effect [108], which blurs borders between tissues, because voxels often represents more tissues in these parts. Modified version of region growing, able to remove this effect, was introduced by Lakare [63]. It was later shown in Salman's study [105] that this method improved accuracy of brain tumor segmentation in 3D T1 MR images.

Region growing was used in several works as a refinement step, e.g. Dou et al. [30] used it as an adjustment of fuzzy segmentation in multisequence MRI. Another example is the work of Rexilius et al. [98]. They used progressive region growing to refine results of a segmentation algorithm based on histogram analysis of multisequence images, namely T1C, T2 and FLAIR. This was carried out by histogram matching over generated probabilistic models. However, generalization of their method was difficult due to the training of the histogram model for fully-enhanced tumor. Moreover, an intensive user interaction was desired in their

method. Kumar and Halder [61] used region growing for the brain tumor segmentation task after applying dynamic clustering and detection of the cluster with the highest intensity, which approximately determined the brain tumor location. The seed point was then examined by an asymmetry map. Zhang [130] also used region growing as a refinement of the segmentation results, where multisequence brain tumor segmentation using multi-kernel SVM was performed.

2.2.2 Watershed

Another region-based method, which was used by several researchers for brain tumor segmentation, is watershed. This term refers to a ridge that divides areas drained by different river systems. A catchment basin is the geographical area draining into a river or reservoir. The image is considered to be a landscape where voxel intensities represent the altitude and each valley, i.e. low intensities, is associated with a catchment basin. All basins are continuously flooded and dams are built in places where water coming from different basins meets. The process finishes when the highest point, i.e. highest intensity, is reached.

This technique was used for brain tumor segmentation by Letteber et al. [70], and Dam et al. [24]. Both methods were based on multi-scale watershed segmentation. However, both methods were semi-automatic and required a lot of user interaction to achieve accurate results. The user had to select or deselect created segments in order to sculpt the desired object. Another method using interactive watershed segmentation was introduced by Mancas and Gosselin [75].

Cates et al. [13] carried out a study evaluating the effectiveness of a three-dimensional interactive image segmentation technique relying on watershed. They showed improvement in subject interaction times and inter-subject precision over hand contouring. Their analysis identified failures in places where edges were poorly defined in the data and raised questions about the wisdom of using expert segmentations to define ground truth.

The problem of watershed lies in over-segmentation. Kong et al. [57]introduced merging process for over-segmented brain tumor MR images using fuzzy c-means. Another region merging algorithm for watershed segmentation outputs was proposed by Bleau and Leon [9] and applied to medical images. Different approach was applied by Kaleem et al. [52], who pre-processed input images to avoid over-segmentation instead of post-processing based on region merging. This pre-processing consisted of noise reduction, attenuation of the curvilinear structures present in MR images, and computation of morphological gradients and markers of the input image. However, the presented results were not satisfying. Similar pre-processing step with better results was applied by Singhai and Ladhake [113].

2.3 Edge-based methods

Edge- or boundary-based methods overcome the limitations of region-based methods by searching for boundaries between different tissues. They use gradient information in local or global areas searching for high gradient values usually representing an object or a tissue boundary. These methods require initial estimation of the contour, which may be done interactively or by another segmentation technique. This contour is then refined according to the gradient information in a local area and the algorithm parameters keeping the contour smooth. These methods are based on deformable models, which can be divided into two groups: parametric and geometric. Both of them formulate a propagating interface (a curve in 2D and a surface in 3D) which moves with a certain speed.

2.3.1 Parametric deformable models

Parametric deformable models are also known as *snakes* or *active contours*. These methods move a deformable contour C through the image spatial domain minimizing an energy function $E(C)$ given by the equation:

$$E(C) = E_{int}(C) + E_{ext}(C), \tag{2.1}$$

where $E_{int}(C)$ express an energy function dependent on the regularity of the curve, while $E_{int}(C)$ is an energy function dependent on the image information. Both energy functions are known as internal and external forces [77].

Khotanlou et al. [56] used parametric deformable models as a refinement step for segmentation of various types of brain tumors in T1 images. Disadvantage of methods based on parametric contours is the requirement of a good initialization, otherwise the model have to be re-parametrized during the segmentation process. Classical snakes based on gradient information have also problem in areas without clear boundary. Another drawback of these methods is their capability of dealing with topological changes such as merging and splitting.

Several attempts have been made in order to improve snakes' performance. Law et al. [64] sub-sampled the processed boundary points, and, hence, reduced the computing time. They used this methodology for 2D brain tumor segmentation. Luo et al. [74] combined Gradient Vector Flow snakes and the balloon model for brain tumor segmentation in 2D T1 images. The work was then extended by the same authors into 3D in [73].

2.3.2 Geometric deformable models

Geometric deformable models are also known as *level sets*. They were proposed by Caslles et al. [12] to overcome limitations of active contours, especially the topological adaptation. Since geometric deformable models are independent of the parametrization, the contour (or surface) can be represented as a level set of a high-dimensional function. Therefore, changes in topology can be handled automatically [128]. The level set ϕ evolves with speed $v(k)$, where k is a curvature, along its normal direction according to the equation:

$$\frac{\partial \phi}{\partial t} = V(k) \left| \nabla \phi \right| . \tag{2.2}$$

Level set techniques have been widely applied to brain and brain tumor segmentation in MR images as well as to the whole medical imaging field. Droske et al. [31] used level sets as a semi-automatic method for brain tumor segmentation in 2D MR images using user's interaction for initialization. This was applied to 3D volumes as independent segmentation of 2D slices and subsequent rendering into 3D. Prastawa et al. [94] used level sets together with the intensity distribution in T1 and T2 images for detection of outliers that were considered to be a tumor part. Cates et al. [14] used geometric deformable models for segmentation of active tumor in 3D T1C volumes. The segmentation process started with an interactive determination of a tumor region by the user in one or several slices. Ho et al. [48] used a tumor probability map vs. background as information for the level set evolution for segmentation of brain tumors in 3D MR. Lefohn et al. [69] introduced a method for tumor segmentation based on interactive level sets where manual estimation of an initial contour was required. Segmentation of brain tumors by level sets using appearance priors was also proposed by Cobzas et al. [19] and Popuri et al. [92]. Thapaliya et al. [119] proposed a level set method for brain tumor segmentation using automatic selective local statistics. Taheri et al. [118] proposed a threshold-based speed function for level sets and used this method for 3D brain tumor segmentation. The selection of the threshold was related to the convergence rate and the segmentation accuracy and it was determined by the curvature k.

2.4 Atlas-based methods

Atlas is a widely used prior knowledge in the brain tumor segmentation domain. It is able to determine the spatial or intensity distribution of a given structure. The atlas is created manually. The use of the atlas requires application of a co-registration technique in order to align the atlas to the segmented image or volume and the

quality of the segmentation output depends on the quality of the co-registration process.

The inter-patient atlas, which consists of a spatial probabilistic distribution of three tissues that are present in healthy brains, is the most widely used atlas in brain tumor segmentation. Such atlas is depicted in Fig. 2.1. Methods based on such atlas usually searches for the deviations from the atlas, which are then considered to be a potential tumor. This information is commonly used in classification methods that will be described in Sec. 2.5.

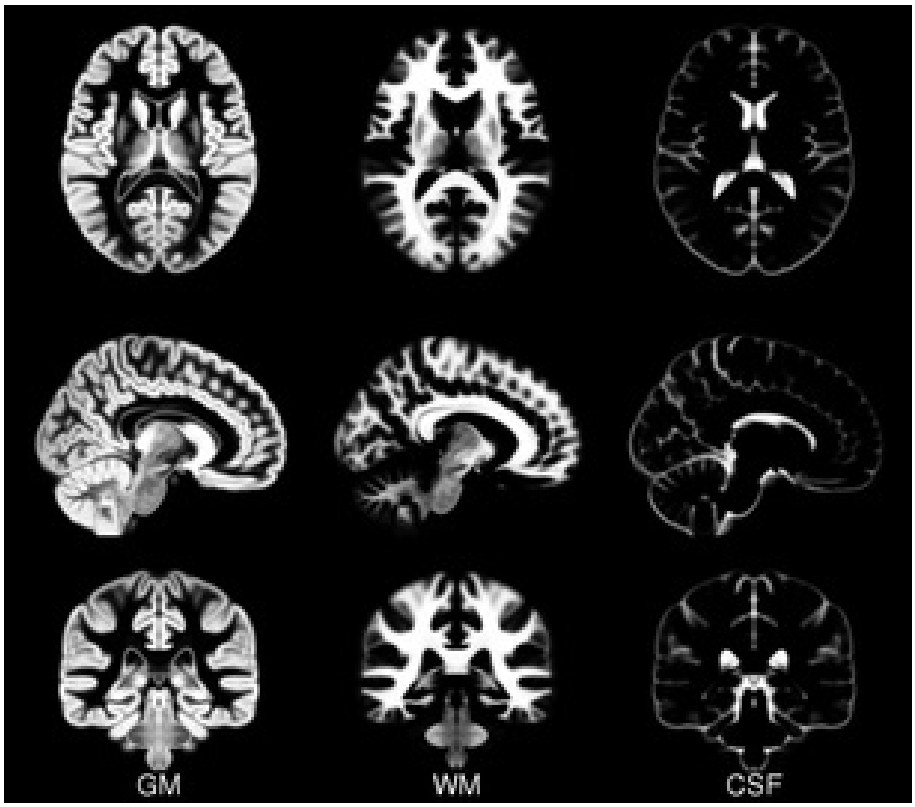


Fig. 2.1: An example of a spatial probabilistic atlas for white matter, grey matter and CSF in axial, sagittal and coronal slices [37].

Moon et al. [84] proposed a method using the spatial probabilistic atlas based on the Expectation Maximization algorithm. Similar method was published by Prastawa et al. [95] where the inputs of the algorithm are T1 and T2 images with prior probabilities from the atlas.

Menze et al. [81] introduced a generative atlas-based approach for segmentation of brain tumor in multisequence data. At first the statistical model of the image formation is built and it is then used in a fully automated segmentation process. The model consists of a normal state, derived as a spatially varying probabilistic shape prior for each healthy tissue, and a tumor state, estimated as a spatially varying

latent probabilistic atlas. Then the probability of a tumor presence is estimated for each channel in multisequence data. Considering the fact that the image intensities of each tissue have the Gaussian distribution in each channel, the joint probability of the latent atlas and the observed variables is computed as a product of both components. The model parameters are estimated using the Maximum Likelihood estimation. The tumor appearance is estimated for each channel in the data.

Another group of atlas-based methods includes methods that use a patient-specific atlas. These methods are broadly used for longitudinal segmentation and modeling of the tumor growth. Kyriacou et al. [62] introduced an algorithm using a biomechanical tumor growth model. The first step of the method is a manual delineation of the brain and the ventricles. Via a uniform contraction model, the tumor is reduced and the simulated prior healthy brain anatomy is estimated. The healthy brain is then registered to the atlas where the estimated tumor growth model is applied. Dawant et al. [26] derived the tumor growth model from optical flow principle and they extended the atlas-based segmentation of the normal brain to the brain tumor segmentation task. This extended algorithm consisted of global and local registration of the atlas to the volume, where the local registration was based on the optical flow principle, with the tumor model estimated from the patient data, and subsequent deformation of the atlas. Menze et al. [82] proposed a combination of a generative tumor growth model, which was based on a reaction-diffusion framework, with longitudinal multisequence data.

2.5 Classification and Clustering methods

This group of algorithms determines the label of every pixel or voxel in supervised (classification) or unsupervised (clustering) manner. Classification algorithms are commonly used in a wide spectrum of medical image processing. They require a set of training data to train a model which can be subsequently applied to previously unseen data. The training data consists of two parts: representation of the data in a feature space (pixel attributes such as intensities, local texture etc.), and labels (manual classification of the data). Clustering algorithms do not require training data and manually created labels. These algorithms automatically search for natural structures and patterns in the data.

2.5.1 *k*-nearest neighbors (*k*-NN)

The k-nearest neighbors algorithm (k-NN) is a simple non-parametric classifier. This algorithm is a type of lazy learning methods. It makes no assumption about the

statistical data structure. The test sample label is derived from its k nearest neighbors from the training set in the feature space. The distance may be measured by e.g. Euclidean, Mahalanobis, or Minkowski distance depending on the implementation. The algorithm has problems with under-represented or noisy data. Another disadvantage is a large storage requirement for training samples since all of them have to be available during the classification process.

The k-NN algorithm was used in the first method which segmented all tumor parts. This interactive method was introduced by Vinitski et al. [125]. A multi-class training data set was defined by the user and a previously unseen pixel was labeled by the most frequent label among the k nearest training samples. This method was very sensitive to the tumor inhomogeneities and noise.

Kaus et al. [53] proposed a method based on k-NN for segmentation of meningiomas and low grade gliomas. They used an atlas of approximately 250 brain structures created by medical experts as additional information to the pixel intensity. All pixels were classified into five groups (background, skin/skull, brain, ventricles and tumor), where the homogeneity of the tumor region was expected. Therefore, the method was only applicable to the above mentioned tumor types. Registration to the manually created atlas was employed to resolve mis-classification problems caused by overlap of intensities of different tissues. Khalid et al. [55] used k-NN algorithm for segmentation of abnormalities in FLAIR images of brain. They used 10 nearest neighbors, which are ranked based on the minimum square distance.

Havaei et al. [47] proposed an interactive method for brain tumor segmentation, which uses k-NN followed by conditional random fields (CRF). The method required a rough manual selection of segmented tissue in two slices of the segmented volume. The algorithm was able to segment the volume in one minute after the manual delineation.

2.5.2 Bayesian approach

Methods, which use Bayesian approach, assume a multi-variate Gaussian distribution of pixel or voxel intensities in every tissue. Both parameters of the distribution, i.e. mean and covariance, are estimated from training samples. The name of these methods is derived from the Bayes rule, which they use. They maximize *a posteriori* probability $Pr(c|x)$, where c denotes the class of the observed data x, using the Bayes rule $Pr(c|x) = \frac{Pr(x|c)Pr(c)}{Pr(x)}$, where $Pr(x|c)$ is called likelihood, $Pr(c)$ is *a priori* probability of a class c computed from the training data, and $Pr(x)$ is the probability of the observed value x.

Corso et al. [21] introduced a multilevel segmentation method for GBM tumors based on segmentation by weighted aggregation introduced by Sharon et al. [110].

The segmentation is a unification of a classification based on a Bayesian model and a graph-based affinity calculation. Ain et al. [1] used Bayesian classification for segmentation of tumor core in multisequence MR data. They used Proton Density (PD), T1, T2 and T2* images. The method consisted of three steps: extraction of features using discrete cosine transform, classification of pixels into healthy and potential tumorous groups by Bayesian classifier, and the final extraction of the tumor using the k-means clustering algorithm and the canny edge detector.

Wang et al. [126] proposed a method combining Normalized Gaussian Bayesian classification with 3D Fluid Vector Flow. They model healthy brain tissues (white matter, gray matter and CSF) by Normalized Gaussian Mixture Model, which is then used to estimation of Gaussian Bayesian Brain Map. Candidate tumor region is estimated from this map and it is used as initialization of 3D Fluid Vector Flow algorithm. The final tumor region is acquired after basic post-processing steps such as thresholding and morphological operation.

The problem of above mentioned methods is their assumption of Gaussian distribution of segmented tissue, which is not always the case.

2.5.3 Expectation maximization (EM)

The Expectation maximization (EM) algorithm was given its name in 1977 by Dempster et al. [27], who generalized the method that had been used several times under special circumstances. It is an iterative two-step algorithm used to find maximum likelihood parameters of a given statistical model. The name of the algorithm is given by the names of the two steps: expectation and maximization. In the first step, the probability distributions are updated to the posterior probability $Pr(c_i|x_j, \Theta)$ that voxel x_j belongs to a class c_i given model parameters Θ. In the second step, the parameters Θ are re-estimated using these probabilities. These two steps alternate until convergence.

The EM segmentation algorithm was introduced by Moon et al. [84] and Prastawa et al. [95]. Both works were produced by the same group. Both of them proposed a segmentation algorithm that is guided by a spatial probability atlas of healthy tissues, which provides a prior knowledge about the brain structure. A tumor prior probability is generated from the difference between T1C and T1 images. These two probability atlases, along with T1 and T2 images, are passed into the segmentation algorithm. Besides the parameters of the probability distributions, there are also parameters for the bias field in the model. The algorithm consists of three iterative steps: classification of voxels using current probability distributions and the bias field, updating of the bias field estimation, and re-estimation of the probability distributions. The assumption of this method is the presence of fully-enhanced tumor

and not very large deformations. Later, Prastawa et al. [94] extended the described algorithm to make it work with T2 image only. This method is based on detection of outliers, which are detected as abnormal region using registered brain atlas. Subsequently, the presence of edema and tumor is detected within the abnormal region from image intensities, and, finally, the spatial and geometric constraints are applied to the detected regions.

2.5.4 Support vector machine (SVM)

Support vector machine is (SVM) a supervised machine learning method invented by Cortes and Vapnik [22] in 1995. The aim of the SVM algorithm is to find a hyperplane that best separates the feature space into subspaces according to the training data. SVM is a two-class classification algorithm, however, several extensions have been proposed for multi-class problems [49]. The idea of SVM is to map the data into a higher-dimensional feature space, where it can be linearly separated. The advantage of the SVM method is good generalization of a classification problem. It has become popular in the brain tumor segmentation task and many papers describing this application have been published.

Several simple SVM-based methods have been proposed. Zhang et al. [129] introduced a simple segmentation technique using SVM classification followed by morphological operations for error reduction purpose. Binary segmentation of brain tumor with patient specific training was applied to 2D T1 and T1C images. A similar approach was applied by Mikulka and Gescheidtova [83], who used T2, T1C and diffusion weighted (DW) images for interactive segmentation of brain tumor parts. First, several pixels inside and outside the segmented area are manually labeled. The classifier is then trained and applied to the segmented image. This means that each tumor part has to be segmented separately. This work also described the performance increase in comparison with using only a single sequence.

Garcia and Moreno [39] proposed an algorithm which at first segments brain tumor in 2D slice by Adatron algorithm. The segmentation result is then used as a training set of the SVM algorithm. SVM is used to produce a 3D tumor model. However, the pixel classification task is still performed in 2D. Lee et al. [68] combined SVM with Markov random fields (MRF) to avoid noisy results by incorporating the relationships between neighbor voxels. As the previous methods, they used patient-specific training data. They applied their method to the segmentation of glioblastomas in T1, T1C and T2 images. Later, Bauer et al. [6] employed hierarchical Conditional random fields (CRF) instead of MRF to regularize the segmentation obtained by SVM. They used multisequence intensities and textures as features for voxel classification. Voxels of healthy and tumorous tissues are classified

followed by sub-classification of healthy region into white matter (WM), gray matter (GM), cerebrospinal fluid (CSF), and tumorous region into active tumor, necrotic tumor, and edema.

Another multisequence SVM approach was proposed by Ruen et al. [102], who combined T1, T2, PD and FLAIR MR sequences. The work employed a multiscale scheme, which consists of two steps, to reduce computing time. The aim of the proposed method is not only to segment a brain tumor region but also to track the tumor region during treatment. SVM was also used by Ayachi and Amor [5] for glioma segmentation.

Zhang et al. [131] segmented brain tumors using T2, PD and FLAIR images and used it for evolution quantification during a therapeutic treatment. The SVM algorithm is used in the first step of the method to obtain tumor segmentation. In the second step, the segmentation result is improved using maximum likelihood and distance measures.

2.5.5 Neural networks (NNs)

Neural networks (NNs), sometimes also called Artificial neural networks (ANNs), are statistical learning methods inspired by the human brain. NN consists of a set of simple artificial neurons which are connected together to create a network. These connections are defined by adaptive weights, which are tuned during the learning process. NNs were widely used for classification tasks, however, they were gradually replaced by simpler methods such as SVM due to the computational demands of NNs. Their popularity started to increase again after the introduction of Deep learning approach and its success in several various international image and speech recognition challenges. Since one of the Deep learning methods is used in this work, particularly Convolutional Neural Networks (CNNs), more detailed description will be given in the appropriate chapter.

One of the first attempts to segment brain tumors in MR images using NNs was described by Clarke [18], who used T1, T2 and PD images. Later, Ozkan et al. [87] used NNs for multimodal segmentation of brain images. They used co-registered MR and X-ray images.

Dickson and Thomas [28] introduced a method with two NNs for acoustic neuromas without using patient-specific training. The segmentation is performed by pixel-wise labeling using the intensities of a given pixel and its neighbors. They tested the method on 50 manually labeled images of different patients. The first neural network produces an initial segmentation, which is considered to be a tumor region candidate. The result is refined by morphological operations and passed to the second network.

Reddick et al. [97] proposed a brain segmentation method based on Kohonen self-organizing map (SOM), used for segmentation, followed by a multilayer back-propagation neural network, used for segment classification. The input data consisted of T1, T2 and PD images. Another work using SOM was published by Vijayakumar et al. [123]. Their algorithm segmented tumor, necrosis, edema, cysts and normal tissues from the combination of T2 and FLAIR images. It was also able to specify the tumor grade. Chaplot et al. [15] also used SOM and applied it to the classification of pixels of T2 images into normal and abnornal using wavelet transform for feature extraction. Murugavalli and Rajamani [85] proposed a method combining Hierarchical SOM and Fuzzy C-means to detect various tissues in T1 images.

A lot of NN-based methods started to appear with the increase of the deep learning popularity. Davy et al. [25] implemented a brain tumor segmentation method based on CNNs. The input of the network is a 2D multisequence patch of size 33×33 pixels, which is used to predict the central pixel. The network contains two pathways. The first pathway is a classical CNN with two convolutional layers connected to the whole patch. The second pathway is a fully connected network of fewer layers connected to 5×5 region at the patch center. The latter pathway serves as contour refinement.

Another brain tumor segmentation method based on deep learning was proposed by Urban et al. [122]. Compared to the previous work, 3D patches are used in this algorithm. The network contains three convolutional layers with a typical size of 5×5×5 voxels. These layers contain 8, 16 and 6 convolutional filters, respectively. Pooling layers, which typically alternates with convolutional layers, are not present in the proposed architecture. Convolutional layer with filters of size 1×1×1 voxel is implemented at the end of the whole architecture to speed up the segmentation process. The input of the whole network is a cubic patch of size 9×9×9 voxels, which is used to classify the central pixel. Regions smaller than 3000 voxels are removed during the post-processing step. The training time of the whole network on a single CPU-thread was approximately 30 - 40 hours.

Zikic et al. [137] proposed another brain tumor segmentation algorithm which uses CNNs. The network proposed in this work accepts 2D multisequence patches of size 19×19 as an input, which means that the volume is processed slice by slice. The architecture consists of a convolutional layer with 64 5×5 filters, a max-pooling layer with sub-sampling rate of 3, a convolutional layer with 64 3×3 filters, a fully connected layer and a soft-max layer with five outputs (one for each class).

Recently, Havaei et al. [46] introduced a method that extends the work of Davy et al. [25]. The algorithm uses the same two-pathway CNN architecture but two of these architectures are cascaded, where the output of the first network is used as an

additional feature map for the second network. Since the brain tumor segmentation task is an imbalanced problem, they proposed a two-phase training process. In the first stage, the whole network is trained using a balanced data set. In the second stage, only the output layer is re-trained using the original imbalanced data set. They also proposed bounding the absolute kernel weights and using L1 and L2 regularization together with Dropout method to avoid over-fitting. The method is one of the fastest with the computing time between 25 seconds and 3 minutes.

The common characteristic of all described algorithms is the independent classification of a single pixel, which often leads to noisy results and, hence, post-processing steps are usually required.

All the above described methods based on deep learning used T1, T1C, T2 and FLAIR images from Multimodal Brain Tumor Segmentation Challenge [1], which is also used for evaluation in this work.

2.5.6 Random forest (RF)

Random forest (RF) is an ensemble of decision trees used for classification and regression problems. Decision trees often suffer from overfitting. This problem is reduced in RF by combination of multiple decision trees and their contribution to the final decision. Decision tree (or classification tree) is a simple algorithm, which can be represented by a tree-like graph, where the classification task into a given number of groups using a predefined criterion is performed in each node. This process hierarchically splits the data into smaller subsets. During learning, these criteria are estimated based on the information gain after the split. The tree leaves represent a probability distribution over the classes for a given sample. The final class is determined by combining the distribution of all trees in the forest. All trees are trained randomly using, e.g., random subsets of training samples or features.

The advantage of RF is the possibility to analyze the feature importance since the most important features are most frequently used in the first layers. The less important features can be then removed.

Random forest has been frequently used for the brain tumor segmentation problem as well as for other segmentation tasks in recent years. As described in [41], the advantages of using RF in medical image analysis are the efficiency in dealing with high-dimensional feature spaces, their nature to deal with multi-label problems, and generalization ability. In BRATS challenge in 2013 and 2014, seven out of 17 contributions were based on RF. All methods usually differ in extracted features.

Geremia et al. [40] introduced a method based on spatial decision forests for segmentation of Multiple Sclerosis lesions in 3D MRI. In their algorithm they used T1,

[1] http://braintumorsegmentation.org/

T2 and FLAIR images, knowledge of tissue classes, and long-range spatial context together with symmetry features.

Zikic et al. in [136] and later in [41] segmented brain tumors using the combination of six sequences: T1, T1C, T2 turbo spin echo, FLAIR, and two channels of diffusion tensor imaging. They combined discriminative decision forests with a generative tissue appearance model, which was estimated by Gaussian Mixture Model (GMM). As features, they also used the intensity difference between a given voxel and a voxel with a given offset in the same or a different channel, the same type of difference feature in a cube around the voxels, and the intensity range between these two voxels in one channel.

Festa et al. [35] used 120 features for brain tumor segmentation from four sequences including intensities and the differences between each combination, mean, sum, median and range in 3D neighborhood of four different sizes, the difference between the intensity of a given voxel and the mean value in its neighborhood, edge density and local binary partition from 3D neighborhood and 2D texture features from each plane. During the post-processing step, all regions smaller than seven voxels were labeled by the label of the surrounding area. The computing time of the proposed algorithm is 20-25 minutes per patient running on standard CPU.

Meier et al. [78] made use of a generative version of random forest called density forest, which can be described as hierarchical GMM using hard label assignment. It is used to model class-dependent likelihood $Pr(x|y)$ which estimates the initial tissue probability. The tissue probability estimation is followed by a discriminative classification forest and subsequent regularization using CRF. For the classification of each voxel, they used first-order texture, gradient and symmetry features with probabilities estimated by the density forest summing up to 44 features. In their later work [79], they abandoned the generative density forest and enhanced the feature set instead. It led to the computing time reduction and the performance increase. They extracted context-sensitive features using affine registration of the data with the brain atlas. Particularly, they used so called ray features, which are based on the intensity range in four directions and two distances from the classified voxel, and symmetry features computed as a difference between the intensity of the voxel and its corresponding point in left or right hemisphere. The average computing time of the latter method is 5 minutes per patient.

Texture features were the basis of the work published by Reza and Iftekharuddin [99]. Besides common intensity features, they used a fractal piece-wise triangular prism surface area and texton features. Later in [100], they extended the feature set by a multi-fractional Brownian motion.

Goetz et al. [45] proposed extremely randomized trees, which introduced more randomness during the training. They extracted 54 features from each of four se-

quences, such as intensities, local histograms, first and second order statistics, and histogram based segmentations.

In [41], Geremia et al. provided rough initial class probabilities as another input of RF. They used posterior probabilities which were based on GMM estimation for each tissue. The same features as in [136] were used. Prior et al. [96] implemented RF as a classification technique in their work dealing with active tumor extraction from several MR sequences: T1, T1C, T2, FLAIR, Susceptibility Weighted Imaging (SWI), Apparent Diffusion Coefficient (ADC), TraceW, and rCBV.

Pinto et al. [90] recently published a brain tumor segmentation method using intensity, gradient and context-based features, which are used for pixel classification by RF with 100 trees of maximum depth of 45. The segmentation is followed by morphological operations during the post-processing step.

Common characteristic of all described algorithms is the same as in NN-based segmentation - the independent classification of a single pixel/voxel.

2.5.7 Clustering

Clustering is an unsupervised method of grouping data samples into several groups according to their similarities in the feature space. These algorithms work with current data and do not need a training phase and a training data set. Common methods used for clustering are k-means, fuzzy c-means (FCM - also known as soft k-means), Guassian Mixture Models (GMM), and hierarchical clustering. All these methods, except the hierarchical clustering algorithm, require a specified number of clusters as the additional input to the data. k-means is an iterative algorithm minimizing the sum of distances from each data sample to the centroid it is assigned to. Fuzzy c-means generalizes the k-means method by allowing the percentage of belonging into a given group instead of hard assignment. GMM is a model assuming that the data has been generated by a finite number of normal distributions. The cluster is than assigned by the appropriate Gaussian component in the model. The hierarchical clustering method clusters data by a cluster tree where clusters at one level under a given node create a single cluster in the upper level. All of these methods are sensitive to noise and inhomogeneity.

Phillips et al. [88] used FCM for the segmentation of tissues present in a pathological brain. They combined T1, T2 and PD images. They showed the intensity overlap even with using multiple MR sequences and, hence, the tumor was not segmented accurately. Better results were achieved by Clark et al. [17] and Fletcher-Heath et al. [36] who implemented knowledge-based techniques together with FCM and used them for the same combination of MR sequences. Other authors, such

as Shen et al. [111] or Szilagyi et al. [117], proposed employing dependencies of neighbor pixels in order to overcome the noise sensitivity of FCM.

Zhao et al. [133] proposed a method based on GMM and active contours. GMM served as the initial segmentation for the active contour method, which than refined the boundaries between tissues.

2.5.8 Random fields

Random fields are commonly used as the post-processing step to incorporate neighborhood relations for spatial regularization to correct voxel- or pixel-wise segmentation results. However, they can be used as an unsupervised segmentation technique too. They improve the clustering methods by incorporating the spatial information. Commonly used random fields in image segmentation are Markov random field (MRF) and Conditional random field (CRF).

Random field is a set of random variables in an undirected graph. In case of 3D volumes, random variables and edges represent voxels and relations between neighbors, respectively. MRF is a generative model, which models the joint probability of variables. CRF is a discriminative model, which models the conditional probability of variables.

Capelle et al. [11] proposed a 2D unsupervised method where maximum *a posteriori* (MAP) is used to estimate the realization of MRF. The method segments input brain slices into homogeneous regions and subsequently detects the tumor presence. Gering [43] used MRF as a refinement step after the EM segmentation. His approach used a multi-layer MRF to incorporate a structure in order to constraint the boundary. However, his method is applicable only to homogeneous tumors. Solomon et al. [114] included MRF directly to the EM segmentation and applied it to the segmentation of active tumor parts.

2.6 Tumor detection

Apart from tumor segmentation methods, only several methods for detection of the tumor presence and its approximate location exist. The goal of these methods is not the accurate delineation of the tumor boundary, but only fast decision whether the tumor is present together with its location. Such techniques could be used as an initial estimation of the tumor contour which is necessary for some segmentation methods. Saha et al. [104] located tumors in 2D MR images in axial plane using the fast detection of asymmetry by Bhattacharyya coefficient. The output of the algorithm was the bounding box around the tumor. Ambrosini et al. [3] and Farjam et al. [33] used template-matching technique for this purpose.

3 RELATED WORK - LOCAL STRUCTURE PREDICTION

Medical images show a high correlation between the intensities of nearby voxels and the intensity patterns of different image modalities acquired from the same volume. Patch-based prediction approaches make use of this local correlation and rely on dictionaries with finite sets of image patches. They succeed in a wide range of application such as image denoising, reconstruction, and even the synthesis of image modalities for given applications [50]. Moreover, they were used successfully for image segmentation, predicting the most likely label of voxels in the center of a patch [120]. All of these approaches exploit the redundancy of local image information and similarity of *image features* in nearby pixels or voxels. For most applications the same local similarity is present among the *image labels*, e.g., indicating the extension of underlying anatomical structure. This structure has already been used in medical imaging but only at *global level*, where the shape of the whole segmented structure is considered, e.g. [91, 132].

The first attempt to predict extended 2D patches instead of pixel-wise labels was made by Zhu et al. in [134] and later in [135], who proposed a recursive segmentation approach with recognition templates in multiple layers called Hierarchical Image Models (HIM) for image parsing using structured perceptron learning. This model consists of five layers where the last layer represented 27×27 image patch and each child has only one parent, therefore image patches are non-overlapping. A label patch dictionary is defined manually. Such approach is suitable for approximate delineation of an object in natural image processing but not for accurate segmentation of medical images.

Kontschieder et al. [58] extended the previous work with structured image labeling using random forest. They introduced a novel data splitting function based on random pixel position in a patch, and exploited joint distributions of structured labels. They used image patches of fixed size of 24×24 pixels and label patches of size 11×11. They compared their method to the random forest using common pixel classification in multi-label image segmentation. Later in [59], they compared different sizes of label patches from 1×1 to 13×13 and showed increasing performance with increasing label patch size. For the patch classification, the CIELab raw intensities, first and second order derivatives and HOG-like features were used. For the test function selection, in each node of a tree, the feature patch is associated with a label at a random position in label patch. This random position is generated once per node. The definition of a label patch cluster is not necessary since a random position from the label patch is used for the test function selection in each node of

a tree. The final label for an unseen image patch is then selected as the one with the highest probability in the leaf. Since there are more labels per pixel due to the patch overlap, they proposed a fusion method based on maximizing the agreement between the resulting label patch and the corresponding labeling at a given position, which outperformed simple voting fusion.

Chen et al. [16] used CRF for shape epitome prediction of a given image patch and called their model Shape epitome CRF (SeCRF). The shape epitomes are created using affinity propagation [38] of patches extracted around the image ground truth boundaries. These patch epitomes are represented by a group of shape templates of smaller size, which are generated by translation and rotation of a template mask within the patch epitome. They created a three-level graph model where the first layer represented each patch region encoded by the segmentation template, second level encoded global consistency and the third level encoded pairwise co-occurrence between the second level nodes. In their experiments, they fixed the patch size to 25×25 px. They found that 5 most common patch epitomes contained most of the patches. They also fixed the number of shape template translations and rotations within each shape epitome to 9 and 4, respectively, creating altogether 181 shape templates.

Dollar et al. [29] used the idea from [58] to create an edge detector based on the patch structure prediction using k-means clustering in the label space to generate an edge dictionary, and RF to predict the most likely local edge shape. They used label patches of size 16×16 to be assign to larger 32×32 image patches with stride of 2. They run the prediction for three different scales, the original image and a half and double resolution version.

4 GOALS OF THE THESIS

This chapter describes the objectives and goals of the thesis. The main goal is to develop algorithms for brain tumor detection and segmentation in 2D and 3D single- or multisequence MRI. The proposed methods have to be fully automated, i.e. they have to be able to detect and segment brain tumor without a user interaction. Even though the user interaction is sometimes necessary and even appreciated, e.g. for post-processing correction, the proposed methods must not be based on any kind of user interaction. The proposed methods should be primarily focused on dealing with glioma of both low and high grades. The attention during the algorithms development should be paid to the segmentation accuracy as well as to the computational demands during the segmentation process.

5 METHODOLOGY

The goal of this chapter is to contribute to the medical image analysis domain, specifically to the field of brain tumor detection and segmentation in multisequence MR images. At first, the necessary pre-processing step is be described. This includes skull stripping, intensity normalization, bias field correction and volume registration. In the next part, a fast pathology presence detection approach, which is based on multi-resolution symmetry analysis, is introduced. This method is applicable for 2D slices of brain images of both axial and coronal planes, where the left-right symmetry is usually broken in case of tumor presence. This is followed by introduction of unsupervised extraction of glioma (the most common malignant brain tumor type) from 2D and 3D FLAIR and T2 volumes. This method is based on the previous pathology detection in 2D slices. It uses the same multi-resolution analysis which is then extended into 3D. This is followed by brain tumor locating and subsequent extraction of all brain tumor parts as a whole region. The extraction process is an unsupervised algorithm using the knowledge of glioma manifestation in FLAIR in T2 MR images.

In the last section of this chapter, a supervised approach of three brain tumor segmentation sub-problems (segmentation of whole tumor, tumor core, and active tumor) is proposed. This approach is based on a state-of-the-art Deep learning (DL) algorithm, particularly Convolutional Neural Network (CNN), which is combined with a structure prediction framework where a label structure in a given area is predicted instead of common pixel-wise classification approaches. Pixel-wise segmentation usually suffers from noisy results and requires a sophisticated post-processing step performed, for instance, by Conditional random field (CRF). This step is eliminated in the proposed algorithm thanks to the prediction of labels of more pixels at once based on the image information in the whole region.

5.1 Pre-processing

Most of the segmentation algorithms require pre-processing applied to the input image. In brain MRI processing, this usually includes spatial co-registration, bias field correction, intensity normalization, and skull stripping.

Registration

Image registration is a process of aligning several images in order to have corresponding pixels or voxels at the same position in all images. This is a necessary step in the multisequence image analysis, which allows combining the information

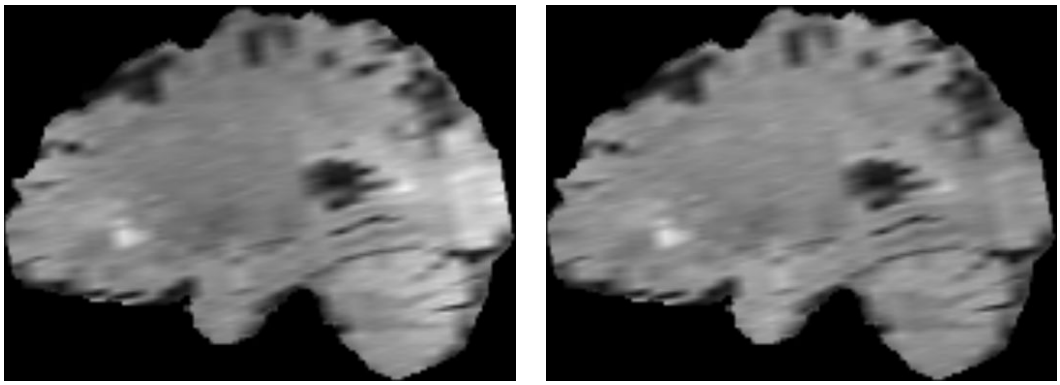
extracted from all images. The goal of the registration process is to find a transformation model that maps the coordinates of a particular pixel from the image to be registered to the coordinates of the corresponding pixel in the reference image.

The data used in this work had been co-registered by a rigid model with the mutual information similarity metric implemented in Insight Segmentation and Registration Toolkit (ITK). Specifically, Versor rigid 3D transformation, where only rotation and translation are allowed, had been applied. As the similarity metric, Mattes mutual information had been used with three multi-resolution levels. The algorithm is based on the work published by Mattes et al. [76]. They proposed using Parzen histograms to estimate the probability density function (PDF). They used zero order and third order B-spline kernel for the reference and registered image intensity PDF, respectively.

Bias field correction

Intensity inhomogeneity is a common problem of MR images caused by patients' position, magnetic settings, and other factors. All classification based segmentation methods use the voxel intensity information where small changes may lead to misclassification. Therefore, the correction of the bias field is required to overcome this problem. The method denoted as N4ITK published by Tustison et al. [121] is used here. This method is an improvement of the nonparametric nonuniform intensity normalization (N3) algorithm. It is an iterative algorithm using B-spline approximation of the data sample distribution, which is based on the least squares fitting technique.

An example of the bias field correction result is depicted in Fig. 5.1 where a comparison between original and corrected images is shown. One can see the intensity decrease in the back part of the brain in the corrected image.



(a) (b)

Fig. 5.1: Bias field correction. (a) original input image, (b) corrected image.

Intensity normalization

Intensity normalization is a critical step for classification based segmentation methods applied to MRI. Two common techniques used for this task are normalization by average intensity and standard deviation, and histogram matching. Both of these approaches work well for images of healthy brains where the shape of the intensity distribution varies only slightly. However, this is not true for images with pathologies, where the shape of the intensity distribution is influenced by the pathology size. On the other hand, as it has been observed, the former approach normalized intensities sufficiently enough for this work's purposes, so the intensities in each volume are normalized so that they have the average value equal to zero and the standard deviation equal to one.

Skull stripping

Another common pre-processing step performed during the brain tumor segmentation process is skull stripping. The aim of this process is to separate the brain tissues from the non-brain intracranial tissues and skull. Several methods developed for this purpose were evaluated by Fennema-Notestine et al. [34]. Since the data used for testing the proposed algorithms had already been skull stripped, this step is not carried out here.

5.2 Brain tumor presence detection

The method proposed for brain tumor presence detection is a symmetry based algorithm for the detection of a brain tumor presence in 2D MR slice. It is a modification and extension of a method that I have already published in [152, 154]. The algorithm uses the prior knowledge of structural differences between healthy and pathological brains. This knowledge is based on the fact that brain tumors break the approximate tissue structure symmetry in the left and right hemisphere, which is usual for healthy brains, unless they are located symmetrically in the middle of the brain, which is not common. Even in case of such tumor location, the symmetry is often broken because tumors usually don't grow symmetrically.

The symmetry prior requires another pre-processing step: mid-sagittal plane detection. This has been studied in several works, e.g. [72, 103]. Another approach uses registration of the brain volume to a reference aligned volume, which may be based on the same methodology described in the section about pre-processing. The proposed method is not sensitive to small deviations from perfect alignment and is able to work correctly even with slight rotation.

The proposed method was designed for both planes where the left-right symmetry can be observed, i.e. axial and coronal. It is a supervised classification algorithm, which uses features extracted from multiresolution asymmetry maps and, in case of application to particular slices in 3D volume, input slices. The latter type of features, based on the input image, are not applicable for stand-alone slices since the intensities in such slices varies due to different measurement parameters and equipment, and the intensity normalization is not possible. Random forest (RF) is used as a classification algorithm but another supervised algorithm may be used.

In this section, a brief overview of the RF algorithm is given followed by the description of multiresolution symmetry analysis and feature extraction.

5.2.1 Random forest (RF)

The Random forest (RF) algorithm has been widely used in medical imaging and brain tumor segmentation as described in 2. A brief introduction into RF is given here, more detailed information can be found in [23] RF is an ensemble of decision trees used for classification and regression problems. Decision trees often suffer from overfitting. This problem is reduced in RF by combination of multiple decision trees and their contribution to the final decision.

Decision tree (or classification tree) is a simple algorithm, which can be represented by a tree-like graph (Fig. 5.2), where a classification task into a given number of groups using a predefined criterion is performed in each node. This process hierarchically splits the data into smaller subsets. During learning, these criteria are estimated by maximizing the information gain I in a given node, which is defined as the uncertainty reduction after splitting the training data. Commonly, the following equation is used:

$$I = H(S) - \sum_{i \in L,R} \frac{|S^i|}{|S|} H(S^i), \tag{5.1}$$

where H stands for the entropy and S indicates the input data set, which is split by a given node into two subsets S^L and S^R. $|\cdot|$ denotes the cardinality for set arguments.

A tree leaf represents a probability distribution over the classes for a given sample. The final class is determined by combining the distribution of all trees in the forest. All trees are trained randomly using, e.g., random subsets of training samples or features.

The advantage of RF is the possibility to analyze the feature importance since the most important are most frequently used in the first layers. The less important features can be then removed.

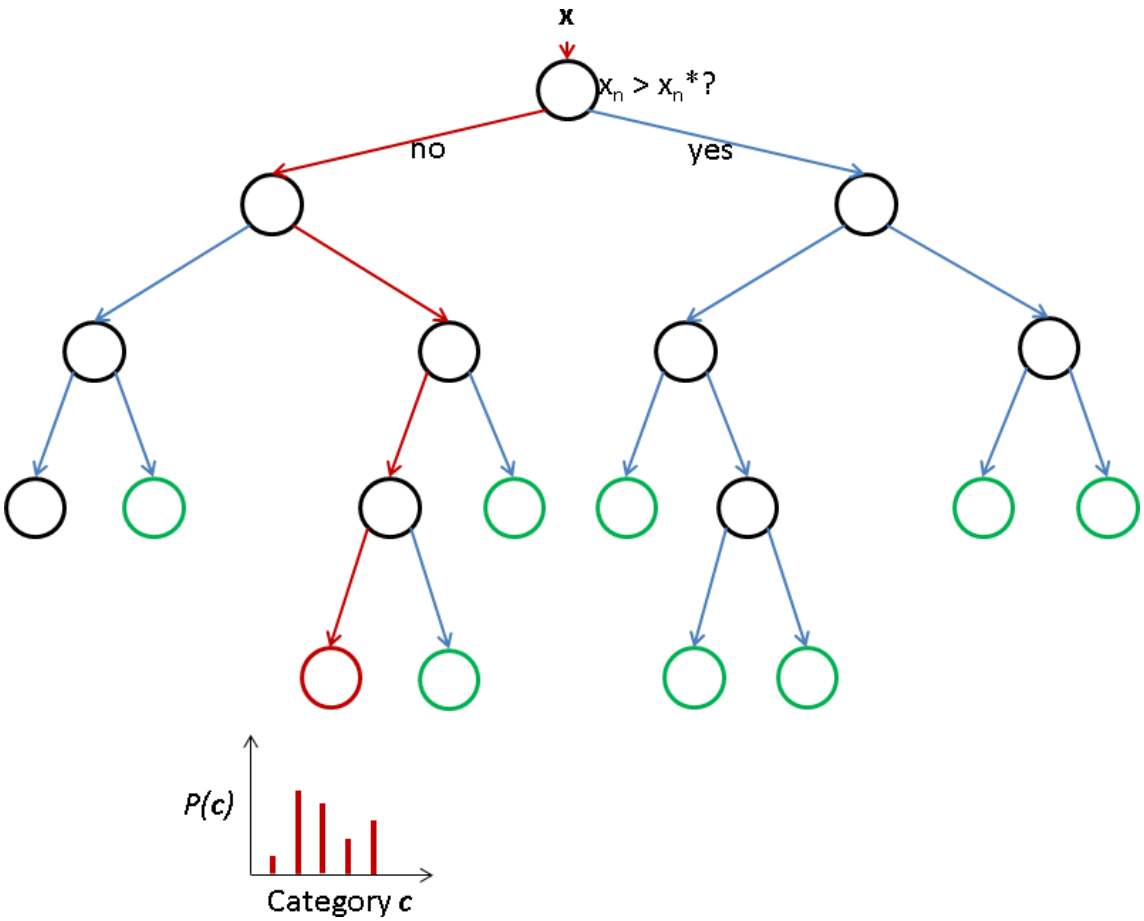


Fig. 5.2: Toy example of a decision tree. Red arrows show propagation of an input sample **x**. Green circles stand for tree leaves, and the red circle depicts the leaf reached by the input sample with given probability distribution depicted bellow.

5.2.2 Symmetry analysis

The first part of the method is the detection of symmetry anomalies, which are often caused by a brain tumor, whose detection is the main purpose of this algorithm. Assuming that the head is aligned, the brain boundary is approximately symmetric, and the image is cropped appropriately, the symmetry axis is parallel to the sagittal plane and divides the image into two parts of the same size. Since the method is not pixel-based, the precision of the determined symmetry axis is not critical.

A square block, with the side length of 10 voxels is created. This size used in multiresolution approach is suitable for the detection of both small and large tumors. The algorithm slides through both halves symmetrically by this block. The step size is smaller than the block size to ensure the overlap of particular areas that are compared with its opposite symmetric part.

The comparison is carried out by the Bhattacharyya coefficient [7], which was also used in [104]. Normalized histograms with the same range are computed from both parts and the Bhattacharyya coefficient (BC) is computed from the histograms as follows [7]:

$$BC = \sum_{i=1}^{B} \sqrt{l\,(i) \cdot r\,(i)}, \tag{5.2}$$

where B denotes the number of bins in the histograms, l and r denote histograms of blocks in the left and right hemisphere, respectively.

The range of values of BC is $\langle 0, 1 \rangle$, where smaller the value, the bigger the difference between histograms. For the subsequent purposes, an Asymmetry coefficient (AC) is computed as $AC = 1 - BC$.

The Asymmetry Coefficient (AC) is computed for all blocks. Since the step size is smaller than the block size, the overlap exists and more values of AC are present for most pixels. To obtain the appropriate asymmetry map, the mean of all values for a particular pixel is computed.

The computed ACs create an asymmetry map where the higher value expresses the higher probability of the tumor presence in particular location. Examples of such asymmetry maps for different image resolutions are depicted in Fig. 5.3.

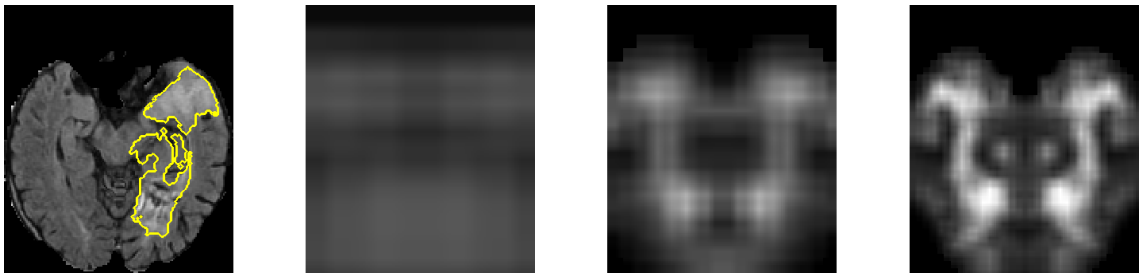


Fig. 5.3: Example of a brain tumor asymmetry map for three different resolutions. The input image is depicted on the left followed by the asymmetry maps with increasing image resolution.

Multiresolution approach

To make the whole process of the symmetry analysis robust, the image is processed in the same way in several different resolutions with constant block size. The input image is iteratively sub-sampled by the factor of two. The number of iterations performed during the symmetry analysis was experimentally set to three.

The output of each iteration is a map of anomalies for a given resolution where the higher number denotes higher probability of the anomaly presence. This anomaly

map is both sided and, hence, the healthy regions, where the tumor is present at the opposite side, are also labeled with high probability of a brain anomaly. The product of values corresponding to a particular pixel is computed leading to the final multiresolution asymmetry map. An example of such multiresolution asymmetry map is depicted in Fig. 5.4.

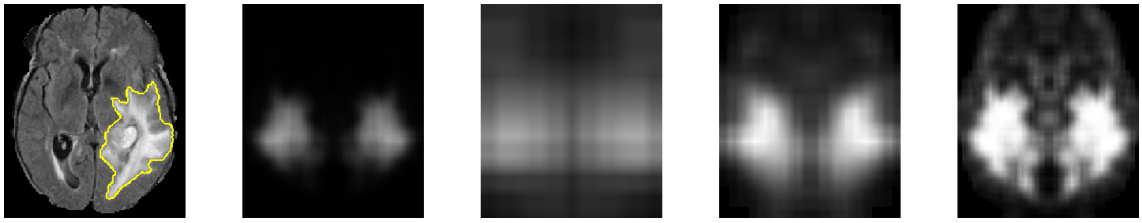


Fig. 5.4: Example of a brain tumor multiresolution asymmetry map. The input image is depicted on the left followed by the multiresolution asymmetry map and corresponding single-resolution asymmetry maps with increasing image resolution.

5.2.3 Feature extraction

When the asymmetry maps are generated, they are used for extraction of features, which are then used for binary image classification into two groups of images showing a healthy or pathological brain. The extracted features are as follows:

- maximum of the multiresolution asymmetry map,
- maximum of each singleresolution asymmetry map,
- number of regions and their size after thresholding the multiresolution map by an absolute value,
- number of regions and their size after thresholding the multiresolution map by a relative value.

Maximum of the multiresolution asymmetry map

Since the proposed method is based on detecting a pathological area by the symmetry analysis, the maximum of asymmetry coefficient is the main feature, which can be used for the image classification. A comparison between multiresolution map of images with healthy brain and brain with tumor derived from T2 slices is depicted in Fig. 5.5 (the second images from the left).

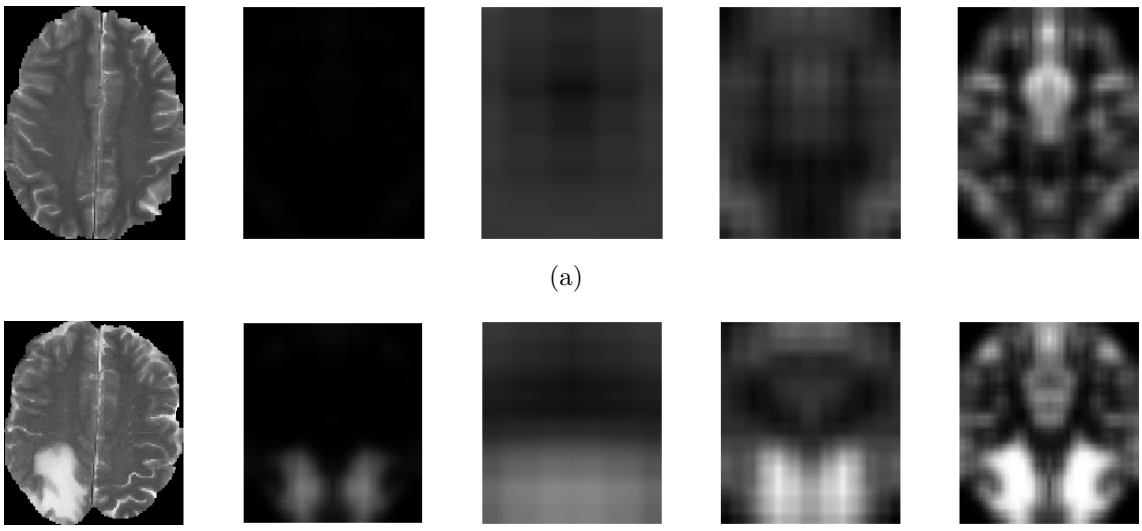
Maximum of each singleresolution asymmetry map

Other suitable features are maxima of each asymmetry map computed in the previous step. Functional dependency of the anomaly coefficient on the image resolution

is non-descending; it means that for lower image resolution, the anomaly coefficient is higher or equal to that for higher image resolution.

For images with large tumors, this value is high even for a low resolution, while for small tumors, this function reaches the maximum later.

A comparison between asymmetry maps of slices with healthy brain and brain with tumor derived from T2 slices is depicted in Fig. 5.5. This figure shows the advantage of the multiresolution approach. For the maximum resolution (the most right image), high asymmetry coefficients exist even for images of healthy brains 5.5(a), here mostly due to the inaccurate alignment. However, they disappear in lower resolutions leading to their suppression in the final multiresolution asymmetry map. On the other hand, as it can be seen in 5.5(b), maximum resolution helps to locate the brain tumor area. Using only lower resolution would lead to vague locating.



(a)

(b)

Fig. 5.5: Comparison between asymmetry maps of (a) healthy brain and (b) brain with tumor derived from T2 slices. From left to right: the original T2 slices, the multiresolution asymmetry maps, and the asymmetry maps for three different single resolutions (increasing from left to right).

A comparison between the maximum asymmetry for multiresolution and each single-resolution asymmetry map in brains with a tumor (green) and healthy brains (blue) are shown in Fig. 5.6.

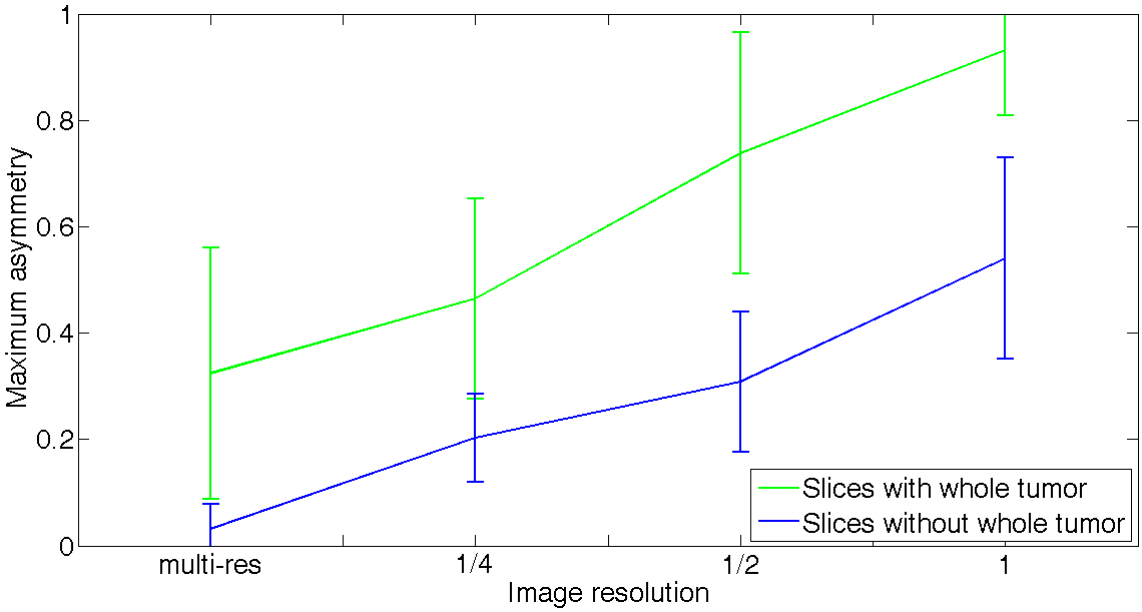


Fig. 5.6: The maximum asymmetry for multiresolution and each single-resolution asymmetry map in FLAIR slices for brains with a tumor (green) and healthy brains (blue).

Number of regions and their size after thresholding the multiresolution map by an absolute value

These features assume that the asymmetry map of healthy brain contains a smaller value compared to the brain with tumor. After the thresholding, in case of healthy brain, there is a smaller number of regions and also a smaller sum of their sizes. In most healthy cases, both numbers are close to zero. Examples of thresholding the asymmetry map by absolute values are shown in Figures 5.7 and 5.8.

Number of regions and their size after thresholding the multiresolution map by a relative value

For the extraction of these features, the multiresolution asymmetry map is thresholded by a relative value computed from the maximum of the map. It is assumed that for brains with a tumor, there is a significant peak in the part where the tumor is situated. Hence, for thresholding by a value derived from maximum, healthy areas are filtered out because they are usually much more symmetric. Moreover, brain tumor is in most cases concentrated in one location, therefore a smaller number of regions is created by thresholding (see Fig. 5.9).

In case of a healthy brain, the property of the feature is opposite. The maximum is comparable to the values in other parts, so more regions are created by thresh-

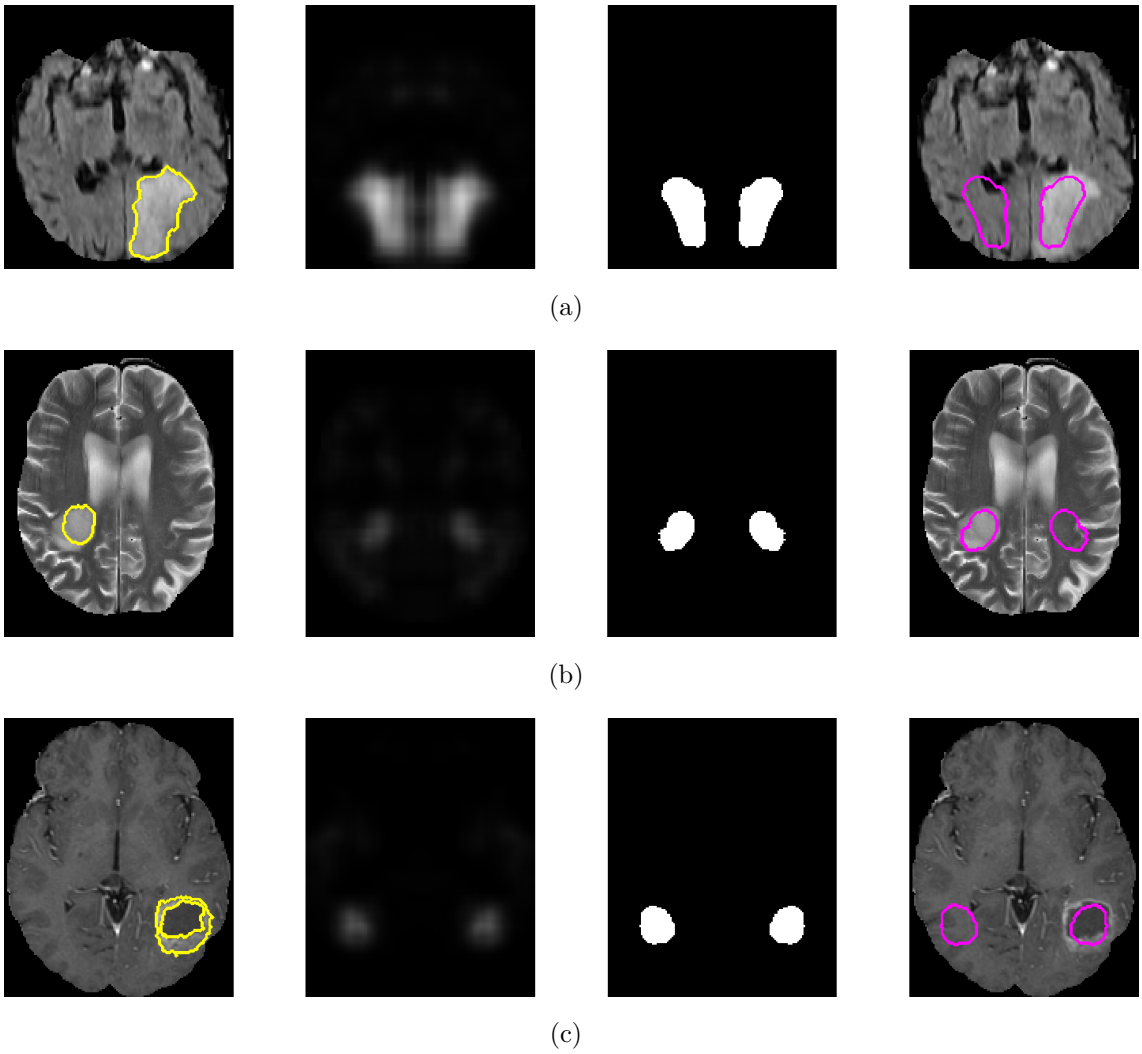


Fig. 5.7: Examples of thresholding multiresolution maps by absolute values. From left to right: input axial slice with manual annotation ((a) whole tumor, (b) tumor core, (c) active tumor), multiresolution asymmetry map, binary mask after thresholding, application of the resulting mask to the input image. (a) input image: FLAIR, threshold: 0.2, (b) input image: T2, threshold: 0.05, (c) input image: T1C, threshold: 0.05.

olding; moreover, they are located in the whole brain. For large tumors, the sum of areas sizes may be comparable to the value computed for healthy brains, but the number of regions is smaller.

Five different levels of threshold are set for both relative and absolute thresholding leading to five values of each feature. Statistical graphs of number of regions and their size for both absolute and relative thresholding for five different threshold levels are shown in Fig. 5.10.

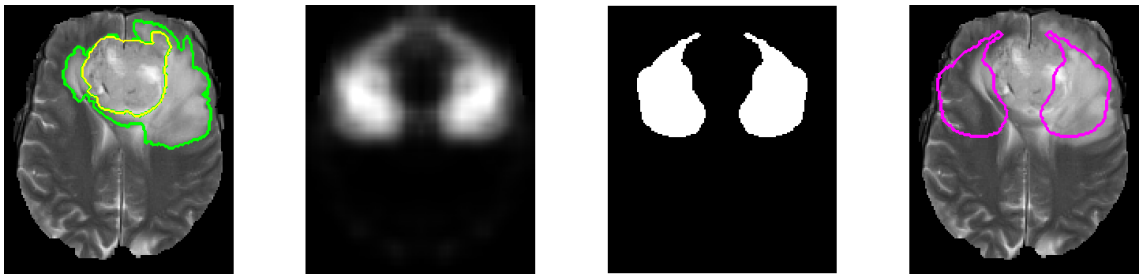
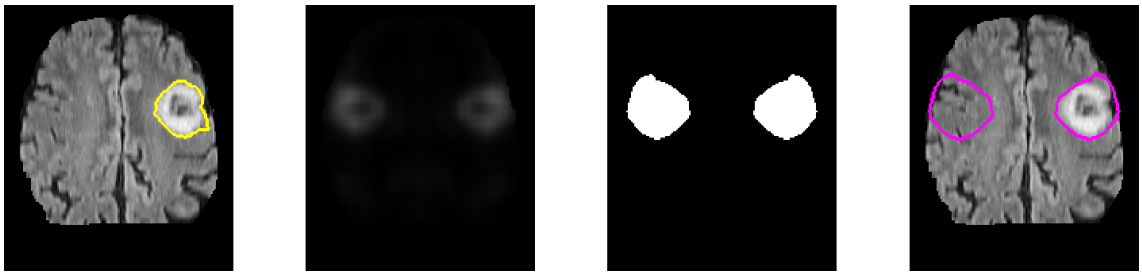


Fig. 5.8: Example of tumor core (yellow) located on the mid-sagittal plane. One can see that the tumor core was not detected correctly using symmetry analysis. However, edema (green), which was not situated symmetrically, was detected.

(a)



(b)

Fig. 5.9: Comparison between multiresolution map relative thresholding (0.2) of (a) healthy brain and (b) brain with tumor derived from FLAIR slices. From left to right: input axial slice with manual annotation, multiresolution asymmetry map, binary mask after thresholding, application of the resulting mask to the input image.

5.2.4 Application to 3D volume

The proposed algorithm may also be applied to particular slices in 3D volume. In this case, additional set of features can be extracted. This feature set assumes that the intensities of the input volume has been normalized. The intensity features are based on the same multiresolution asymmetry map thresholding, but the information is extracted from the input slice. For each threshold level, four more features are extracted from the input image after the application of the mask created by

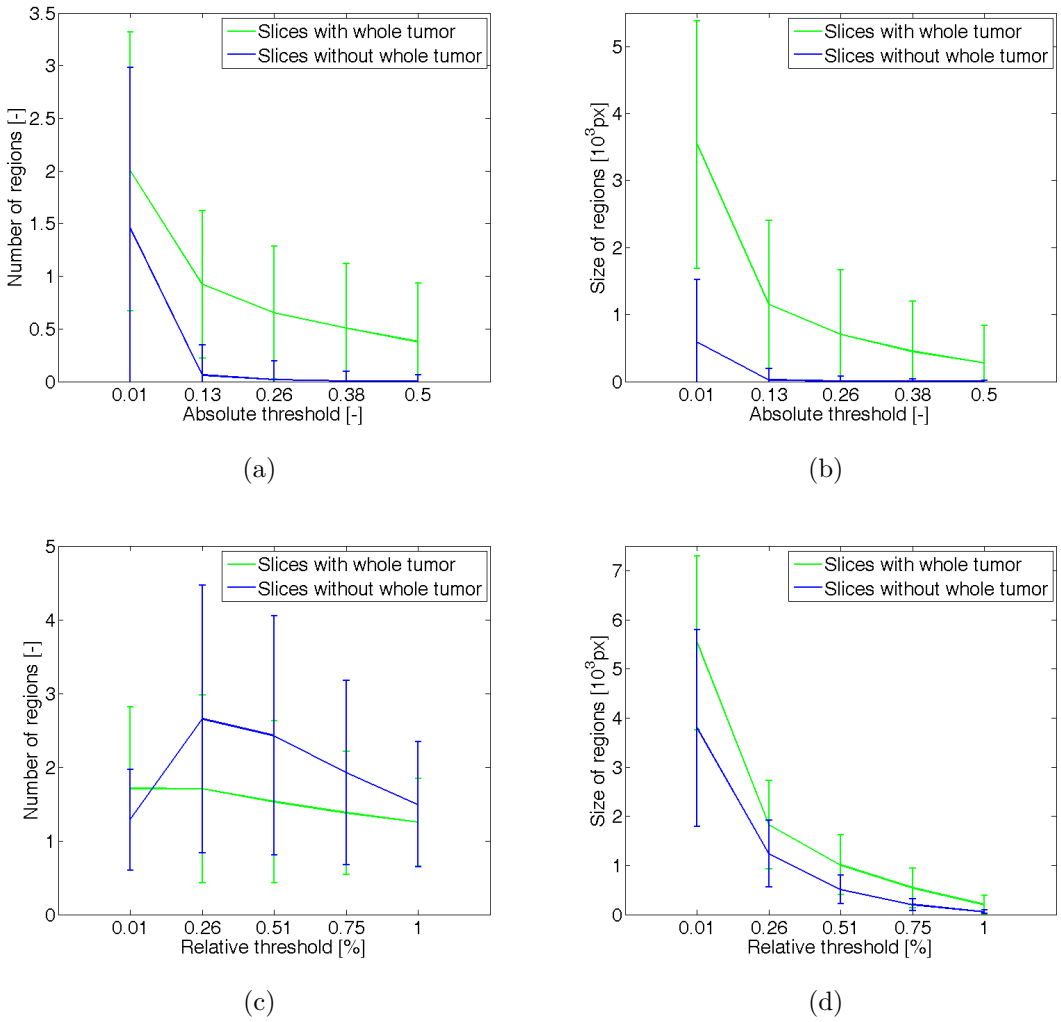


Fig. 5.10: (a) and (b) show number of regions and the sum of their size for different absolute threshold levels of multiresolution asymmetry map derived from FLAIR slices. (c) and (d) show number of regions and the sum of their size for different threshold levels relative to the maximum of a particular multiresolution asymmetry map derived from FLAIR slice.

thresholding. The features are as follows:

- average intensity in the whole masked region,
- absolute value of the average intensity difference in the left and right masked region,
- intensity standard deviation in the whole masked region,
- absolute value of the intensity standard deviation difference in the left and right masked region.

All of these features are computed for both absolute and relative value thresholding. They are depicted in Fig. 5.11 and 5.12 where pathological slices are represented by green color and healthy slices by blue color.

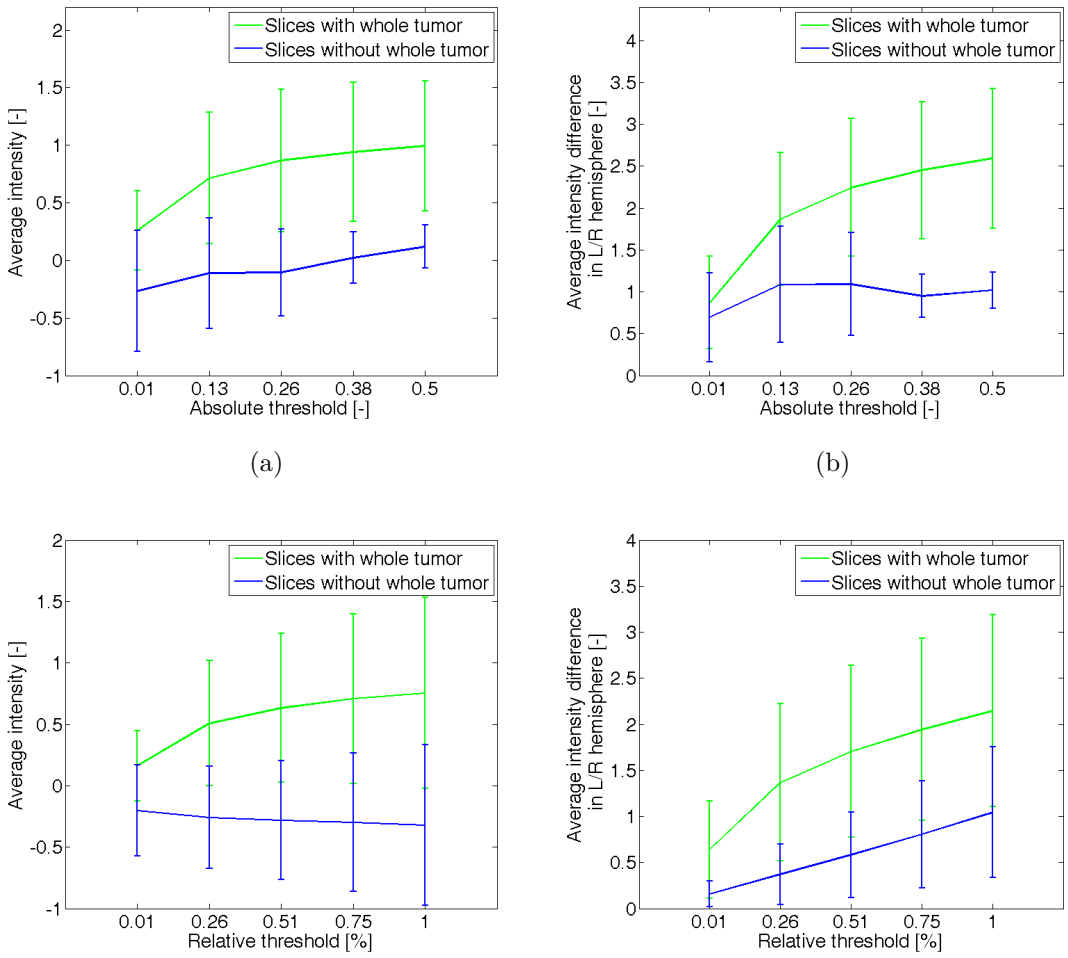


Fig. 5.11: (a) shows the average intensities in the image regions within the mask created by absolute value thresholding of the multiresolution asymmetry map. (b) shows the absolute value of the difference between average intensity in the left and right hemisphere. (c) and (d) show the same values for relative thresholding.

5.3 Unsupervised brain tumor extraction

Gliomas are represented by high intensities in T2 and FLAIR sequences. However, considering different measurement parameters and equipments, the intensities in the resulting image are not known. The common approach is to normalize the intensities

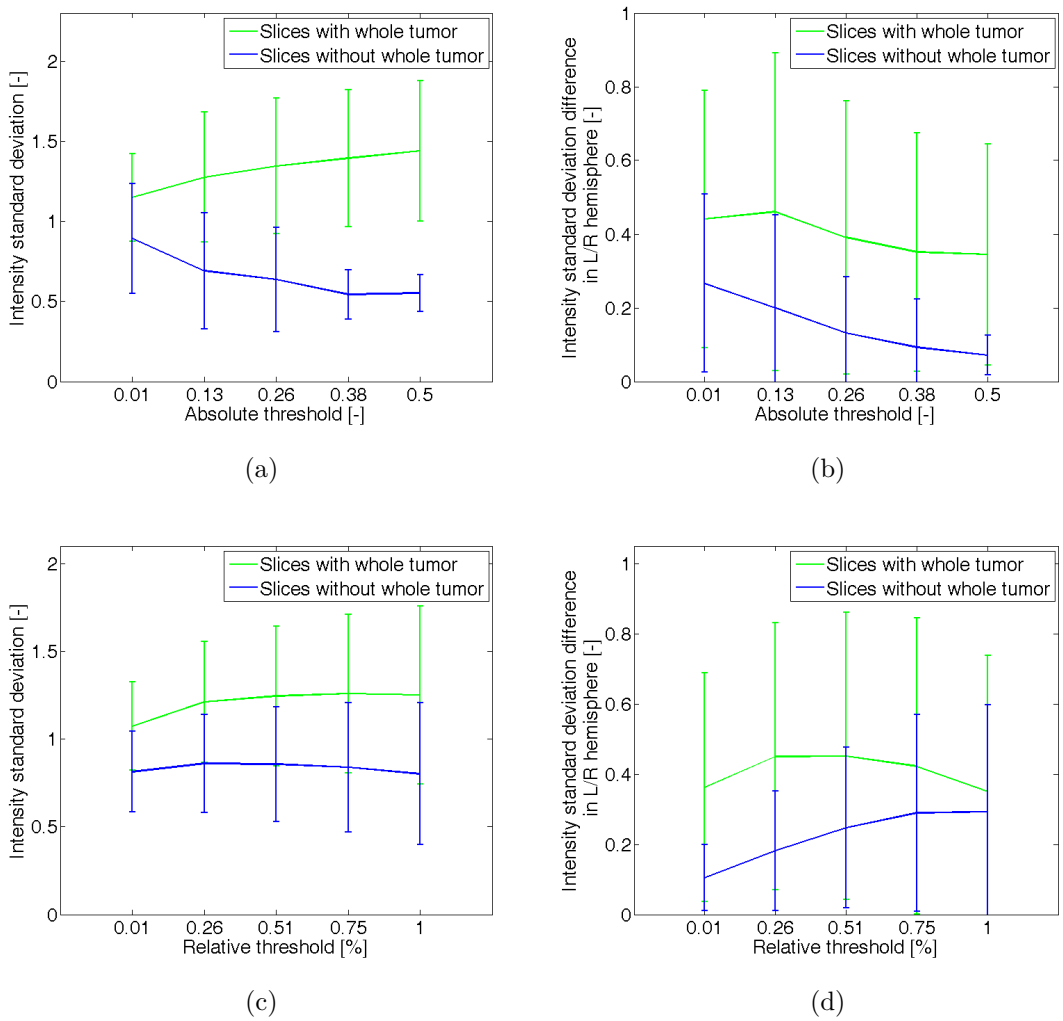


Fig. 5.12: (a) shows the intensity standard deviation in the image regions within the mask created by absolute value thresholding of the multiresolution asymmetry map. (b) shows the absolute value of the difference between regions intensity standard deviation in the left and right hemisphere. (c) and (d) show the same values for relative thresholding.

by standardization or histogram matching algorithm. Considering a stand-alone 2D MR slice, such intensity normalization is impossible due to the lack of information about the tumor size and the slice position. Therefore, machine learning approaches based on the intensity information are not applicable here. The proposed algorithm avoids the intensity normalization by automatic determination of the intensity threshold from the most asymmetric brain parts. It is based on the same symmetry analysis described in Sec. 5.2. Hence, the method requires the same pre-processing as the method proposed for the supervised brain tumor detection. The method can

be applied for both 2D image and 3D volume. Only FLAIR images are used for 2D image application while both sequences are used for 3D volume application. I have already described this method in [146, 148].

5.3.1 Pathology extraction

For the pathology extraction purpose, thresholding of the multi-resolution asymmetry map is performed by the value of 10% of the maximum asymmetry. This value was set experimentally and ensures that at least small region is extracted. The result is a both-sided mask that contains both the tumor on one side and the healthy tissue on the other side.

Since multifocal tumor can appear, the extraction process is not limited to only one region. All regions created by thresholding are considered. As a result, multifocal tumors located asymmetrically can be correctly detected.

The whole area of glioma can be well separated using FLAIR, since they appear hyperintense in this MR sequence. The automatic thresholding can be performed to extract these pathological areas.

The threshold is determined using the Otsu's method [86] from the points inside the resulting mask of asymmetry detection but the thresholding process is applied to the whole image. According to the algorithm proposed by Otsu in 1979, the optimal threshold k^* is a threshold with the following property:

$$\sigma_B^2(k^*) = \max_{1 \leq k \leq L} \sigma_B^2(k), \tag{5.3}$$

where L is the number of intensity values in the region and $\sigma_B^2(k)$ denotes the between-class variance for the threshold k, which is based on class means and is computed according to the equation:

$$\sigma_B^2(k) = \omega_0 \omega_1 (\mu_1 - \mu_0)^2, \tag{5.4}$$

where ω_0, ω_1 mean the probabilities of the class occurrence and μ_0, μ_1 stand for the class mean levels for classes 0 and 1, respectively.

Morphological erosion and dilation are applied to the resulting mask to smooth the region borders and separate regions connected by a thin area. The conjunction of these masks is then found. Since some incorrect areas could be extracted, only those, which are situated mostly inside the asymmetric region, are labeled as pathological. Regions with the size smaller than 10% of the largest segment are also eliminated. Since the pathological area may extend beyond the asymmetry area border, the whole region created by the thresholding is extracted.

5.3.2 Extension into 3D

The same algorithm may be applied to 3D volumes. The multi-resolution asymmetry map is computed in the exactly same way but in 3D instead of 2D, i.e. cubic blocks are used instead of square blocks. An example of a 3D single-resolution asymmetry map is shown in Fig. 5.13. Figure 5.14 shows an example of a 3D multi-resolution asymmetry map. I have already described this 3D extension in [142, 149].

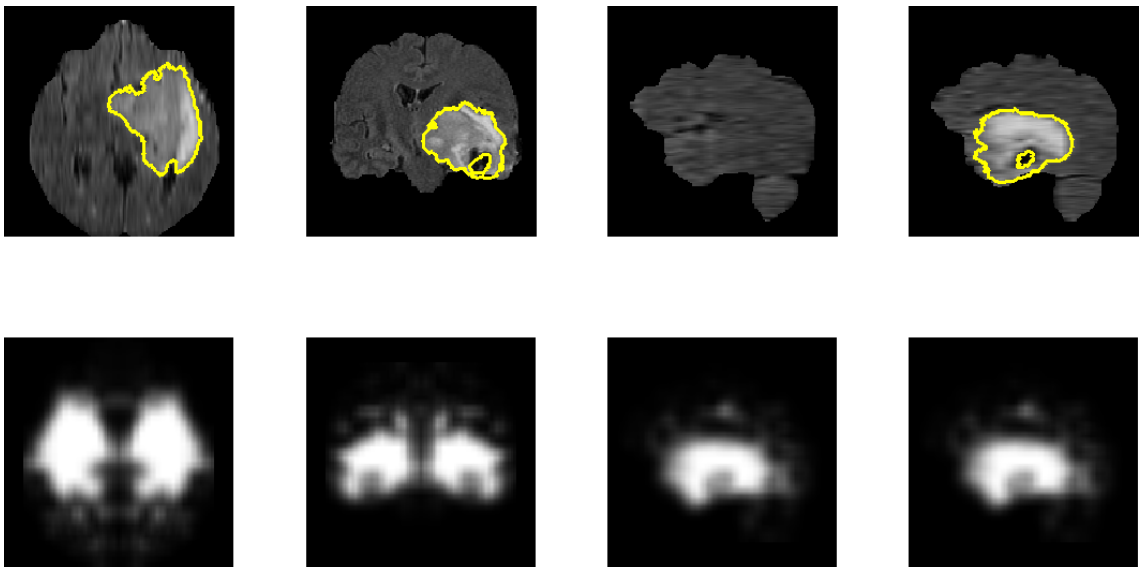


Fig. 5.13: Example of the brain tumor asymmetry map on real data of high-grade glioma for the maximum image resolution. The upper row shows FLAIR slices with the highest AC with ground truth manual annotations. The lower row shows the 3D asymmetry maps in the corresponding slices.

The extraction process starts in the axial slice where the highest AC was detected and it is then propagated into the whole 3D volume. Such approach is more accurate than the extraction directly from the 3D asymmetry map. However, it is slightly slower.

For 3D extraction purpose, T2 volume is used together with FLAIR to improve the accuracy. Since brain tumor appears hyper-intense in both T2 and FLAIR images, the same algorithm is applied to both volumes and their intersection is found. The advantage of the combination of both sequences is depicted in Fig. 5.15. An example of the computed threshold from asymmetric region in T2-weighted and FLAIR image of the same slice is depicted in Fig. 5.15(a), where histograms of the whole asymmetric area (blue bar graph), the true pathological area (green curve) and the computed thresholds (red line) are shown. The combination of both images and their thresholding is shown by a scatter graph in Fig. 5.15(b). As it can be seen

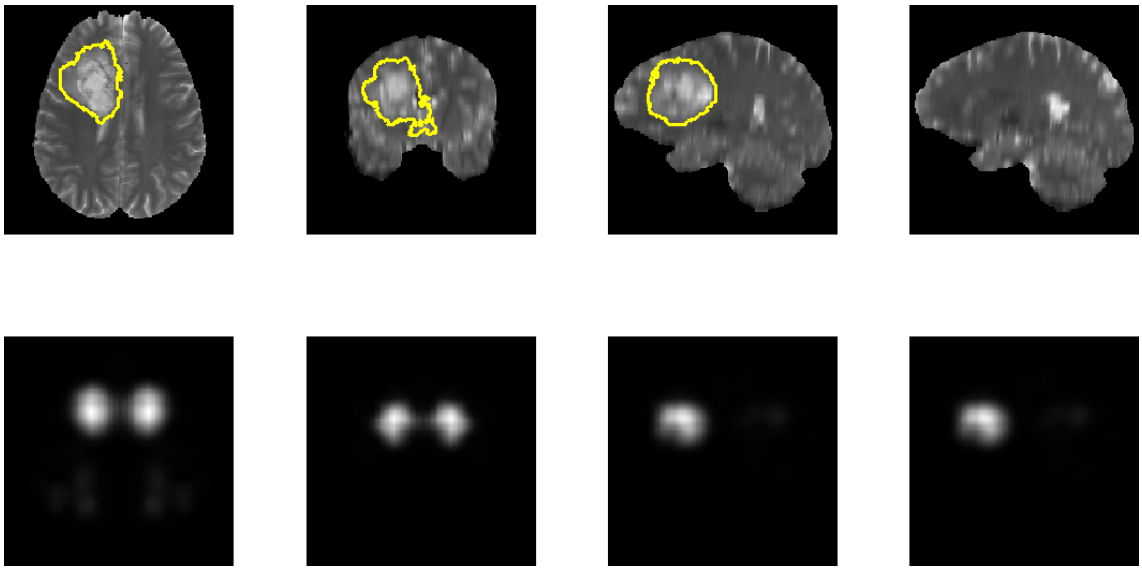


Fig. 5.14: Example of a brain tumor probability map on volumes from real data of the low grade glioma. The upper row shows T2 slices with the highest AC and corresponding manual annotations. The lower row shows the asymmetry maps of the corresponding slices.

in both these figures, the pathology could not be separated properly in neither of them, but combination of both brings improvement.

The resulting mask $\mathbf{M}$ of the extraction process is created as $\mathbf{M} = \mathbf{M}_{T2} \bigcap \mathbf{M}_{FLAIR}$, where $\mathbf{M}_{T2}$ and $\mathbf{M}_{FLAIR}$ denote the thresholding mask of T2-weighted and FLAIR image, respectively. This corresponds to the pixels with intensities situated in top right rectangle in Fig. 5.15(b).

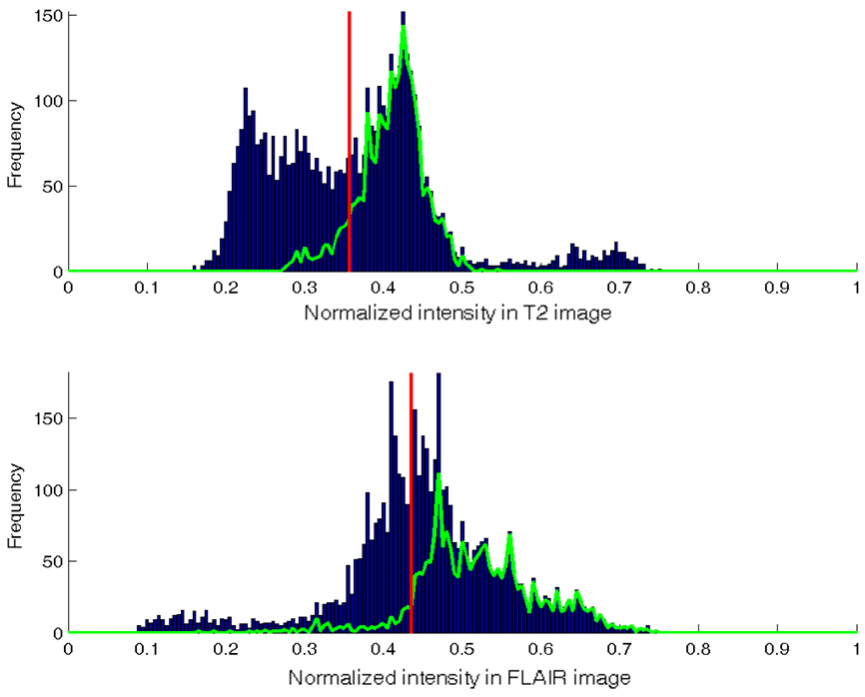
Propagation into neighbor slices

Once the pathology is extracted from the axial slice with the highest asymmetry coefficient, it can be propagated into other slices. At first both 3D volumes are thresholded using the particular threshold values determined in the initial slice with the highest asymmetry. In order to avoid extraction of healthy areas far from the pathological ones, the propagation of mask estimated in neighbor slice is necessary.

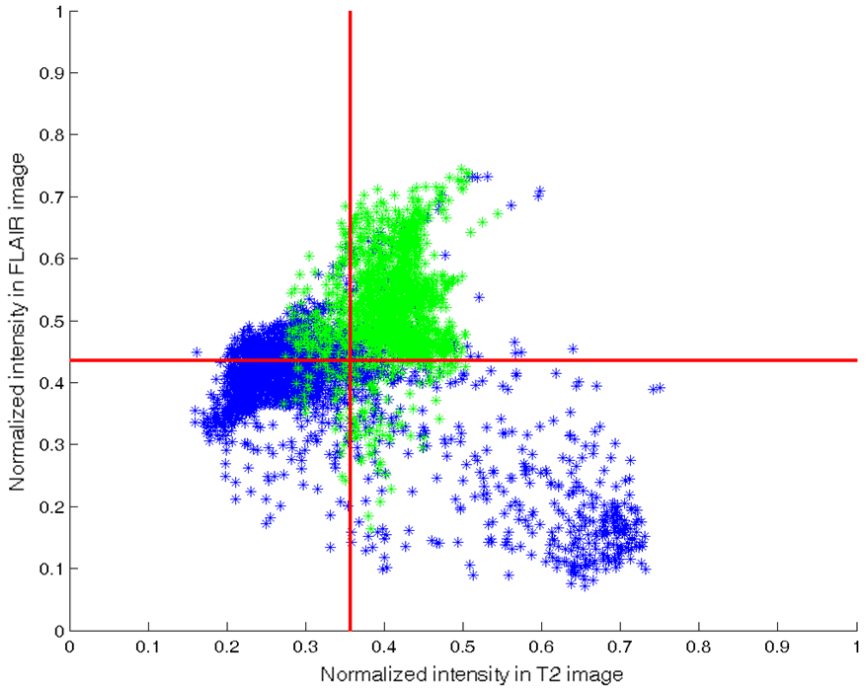
The propagation process starts with both neighbor axial slices and continues in both directions. The result for n-th slice can be defined as:

$$\mathbf{M}(n) = \mathbf{M}_{T2}(n) \bigcap \mathbf{M}_{FLAIR}(n) \bigcap \mathbf{M}_{D}(n \pm 1), \tag{5.5}$$

where $\mathbf{M}_{D}(n \pm 1)$ is the dilated mask from neighbor slice where the sign depends on the propagation direction.



(a)



(b)

Fig. 5.15: Example of tumorous pixels determination using Otsu's algorithm for T2 and FLAIR image of high grade glioma (a) separately and (b) jointly. Blue: the whole asymmetric area. Green: the pathological area. Red: computed thresholds.

5.4 Brain tumor segmentation using structure prediction

In spite of the success of patch-based labeling in medical image annotation, and the highly repetitive local label structure in many applications, the concept of local structure prediction as described in Chapter 3 has not received attention in the processing of 3D medical image yet.

However, approaches labeling supervoxels rather than voxels has already appeared, e.g. hierarchical segmentation by weighted aggregation extended into 3D by Akselrod-Ballin et al. [2] and later by Corso et al. [20], or spatially adaptive random forests introduced by Geremia et al. [42]. Several structure-based methods has also been used in medical imaging but only at *global level*, where the shape of the whole segmented structure is considered, e.g. [91, 132]. Here, the presented work will focus on *local structure* since global structure is not applicable for objects with various shapes and locations such as brain tumors.

The proposed method transfers the idea of *local structure prediction* [29] using patch-based label dictionaries to the task of dense labels of pathological structures in multisequence 3D volumes. Different from Dollar, convolutional neural networks are used for predicting label patches as they are well suited for dealing with local correlation, also in 3D medical image annotation tasks [71] [93].

The brain tumor segmentation problem consists of three sub-problems: identifying the whole tumor region in a set of multisequence images, the tumor core region, and the active tumor region [80]. All three sub-tasks are processed separately, which changes the multi-class segmentation task into three binary segmentation sub-tasks.

The novelty of this work lies in the principled combination of the deep learning approach together with local structure prediction in the medical image segmentation task.

In this section, introduction to the employed clustering and classification methods is given followed by the description of the proposed method, which has been accepted for publication [155].

5.4.1 Patch clustering

Patch clustering is used for the generation of a label patch dictionary. k-means algorithm is used for this purpose.

k-means

k-means is an iterative clustering algorithm minimizing the sum of distances from each data sample to its assigned centroid. The aim is to partition observed samples into k groups with similar properties, where k has to be chosen before the algorithm starts.

The algorithm usually starts with k random centroids selected from the observations. It is followed by alternating between two steps. In the first step, distances of all data samples to each centroid are computed followed by assigning all data samples to the nearest centroid. In the second step, centroids of each cluster are recomputed. These two steps are repeated until no cluster assignment change occurs, or predefined maximum number of iteration is reached.

The described algorithm often leads to a local optimum. This is usually overcome by several independent clustering with random initial centroids.

5.4.2 Patch classification

Convolutional neural network (CNN) is used as a classification algorithm as it has the advantage of preserving the spatial structure of the input, e.g., 2D grid for images. It is a type of modern Deep learning methods.

Deep Learning (DL)

Deep learning (DL) is a new classification approach based on Neural Networks. It differs from other classification algorithms such as SVM or RF in the way how features are learned. In classical approach, features are extracted before the learning of the classification algorithm starts, and they have to be specified by the algorithm developer. In DL approach, features are learned and specified by the classification algorithm itself based on structures and patterns in the training data. The algorithm searches for important characteristics of the data, which may help in the decision process. One of the aim of DL is to replace hand-crafted features with automatic feature extraction and unsupervised or semi-supervised feature learning algorithms.

In image processing, several DL algorithms have been used, such as Deep Belief Network (DBN), Stacked Auto-encoder (SAE), Deep Boltzmann Machine (DBM), or Convolutional Neural Network (CNN). CNN is used in this work, and hence it will be described in detail.

Neural Network (NN)

Neural network (NN), sometimes also called Artificial neural network (ANN), is a statistical learning method inspired by the human brain. NN consists of a set of

simple artificial neurons which are connected together to create a network. These connections are defined by adaptive weights, which are tuned during the learning process. NNs were widely used for classification tasks, however, they were gradually replaced by simpler methods such as SVM due to the computational demands of NNs. Their popularity started to increase again after the introduction of the deep learning approach and its success in several various international image and speech recognition challenges. A lot of different varieties of NN types exist. Since this work makes use of Convolutional Neural Network, only this type will be described in detail.

Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is one of the DL architectures, which was introduced by LeCun et al. [66] in 1990. They used it for handwritten digit recognition. Except input and output layers, it is comprised of one or more convolutional layers, one or more pooling layers and a fully connected layer. Convolutional and pooling layers are usually in alternating order. A pooling layer is often called subsampling layer since the aim of this layer is to subsample the output of the previous layer in order to reduce variance. The back-propagation algorithm is used for network weights learning. This architecture is designed to take the advantages of a 2D grid input, which is convenient for image processing. In several recent works [93, 122], 3D filters have been proposed for medical image analysis. However, it was shown by Prasoon et al. [93] that it is still too demanding to be efficiently used in today's computers. They proposed a different approach for 3D medical images, which they called a triplanar CNN. They used a 2D network for each orthogonal plane separately, i.e. three independent networks were trained. Roth et al. [101] introduced a so called 2.5D CNN. The method extracts corresponding patches for a given pixel from each orthogonal plane and maps them as separated input feature maps, i.e. input channels. Zikic et al. [137] proposed a method where 4-channel 3D patch of size $d_1 \times d_2 \times d_3 \times 4$ is interpreted as a $4 \cdot d_3$-channel 2D patch of size $d_1 \times d_2 \times d_3 \cdot 4$. In this case, d_3-times more 2D convolutions are performed compared to simple 2D network instead of 3D convolutions.

The input of the network is a 2D image patch of a certain size and a number of channels (e.g. three channels in RGB case), and the output is a probability distribution of patch belonging to each class. An example of a CNN architecture, which was used by Krizhevski et al. [60] for image classification, is depicted in Fig. 5.16.

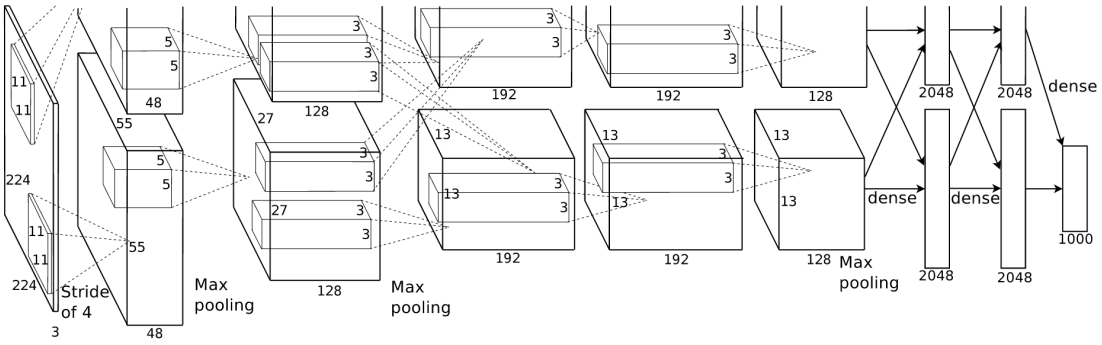


Fig. 5.16: Example of a CNN architecture. This network was used by Krizhevski et al. [60] for image classification.

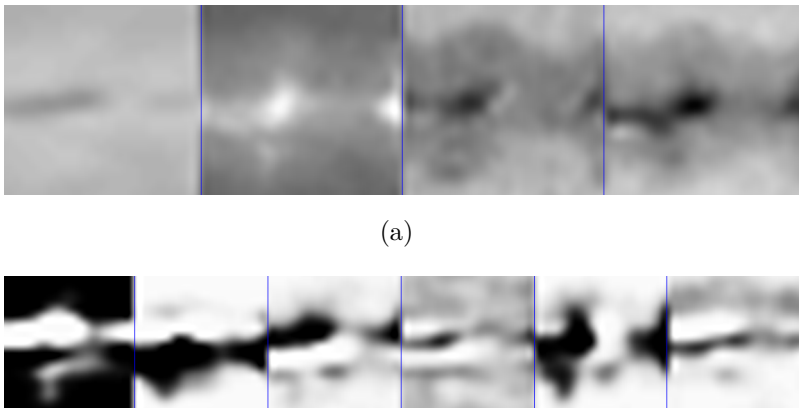
Convolutional layer. The crucial part of the architecture, which also gave the name to this network type, is a convolutional layer [67]. There can be one or more layers of this type. A building block of such layers is a convolutional filter of a predefined kernel size and dimension. Here, standard 2D filters are be used. The number of kernels in a given layer and their size is one of the network parameter, which has to be specified before training. Based on the application, problem complexity and image resolution, this number may significantly vary. Filter kernels are usually randomly initialized and they are tuned during the training phase by the back-propagation algorithm. Outputs of convolutional layers are also known as features or feature maps since convolutional filters are de facto feature extractors. The difference between common feature extractors and feature extractors used in CNN is the fact that they are learned directly from the data instead of being hand-crafted.

The activation of a feature map $\mathbf{o}_j^\ell$, a j-th map in a convolutional layer ℓ, is computed using the equation [10]:

$$\mathbf{o}_j^\ell = f\left(\sum_{i \in M_j^\ell} \mathbf{o}_i^{\ell-1} * \mathbf{k}_{ij}^\ell + b_j^\ell\right), \tag{5.6}$$

where function $f(\cdot)$ denotes a non-linear activation function, which is typically the sigmoid function ($f(x) = (1 + e^{-x})^{-1}$) or the hyperbolic tangent function ($f(x) = (e^{2x} - 1)(e^{2x} + 1)^{-1}$), b_j^ℓ is a bias, M_j^ℓ denotes a selection of feature maps, and $\mathbf{k}_{ij}^\ell$ is a kernel used in the previous layer.

A cascade of alternating convolutional and sub-sampling layers creates a hierarchical cascade of feature extractors where first layer extracts low-level features while the last layer extracts high-level features. High-level features are extracted from lower level features, and only features from highest level are used in the final classification. An example of an input and the corresponding output of a convolutional layer with six filters are depicted in Fig. 5.17.



(a)

(b)

Fig. 5.17: Example of (a) an input (4@24×24 px) and (a) the appropriate output (6@20×20 px) of a convolutional layer with six filters of kernel size 5.

Pooling layer. Pooling, sometimes called sub-sampling, layers are used for sub-sampling of the output of the previous, usually convolutional, layer. This operation makes the network more robust to variance. Early networks used mean-pooling approach, where the average of a receptive field, a non-overlapping squared region used to derive one output value, was propagated to the next layer. In recent years, max-pooling approach, where the average value is replaced by the maximum value, has been introduced. A demonstration of max-pooling process in max-pooling layer with a pooling rate of 3 is shown in Fig. 5.18.

156	150	150	141	138	122	125	125	120
165	155	151	140	135	129	126	124	121
159	152	149	140	128	130	128	127	123
148	145	140	135	131	126	123	125	122
146	142	143	130	130	128	121	120	128
149	140	138	130	129	130	119	118	122
135	135	133	129	134	130	125	128	126
130	135	130	129	135	131	120	130	132
131	130	138	140	138	134	137	131	135

→

165	141	128
149	135	128
138	140	137

Fig. 5.18: Demonstration of the max-pooling process. Picture shows an input and the appropriate output of a max-pooling layer with a pooling rate of 3.

Fully connected layer. The last part of the network, fully connected layer where all neurons of one layer are connected to all neurons of the next layer, is similar to

an ordinary NN. All outputs of the previous layer, either convolutional or pooling, are rearranged to a vector, which behaves as an input to a classical NN architecture. In fact, this is a feature vector extracted from the input patch.

The output $\mathbf{o}$ of the fully connected layer, which is the last network layer L, is defined as [10]:

$$o = f\left(\mathbf{W}^L \mathbf{ov}^{L-1} + b^L\right), \tag{5.7}$$

where $\mathbf{W}^L$ is a matrix of weights, $\mathbf{ov}^{L-1}$ is a vectorized output of the previous layer and $\mathbf{b}^L$ is a bias vector. This part of the network is sometimes called multi-layer perceptron (MLP) since it may consist of more layers.

Back-propagation. Back-propagation, an abbreviation for backward propagation of errors, is a standard algorithm used for weights estimation during the learning phase in NN. It is based on the computation of a cost function gradient with respect to the network weights. The cost function measures the distance between the expected output and the acquired output of the network for all training samples. Back-propagation algorithm for NN and CNN will be briefly described here, more detailed information can be found in [8, 10, 65]. The squared-error cost function E for C classes and N training samples is defined as:

$$E = \frac{1}{2} \sum_{n=1}^{N} \sum_{c=1}^{C} \left(l_c^n - y_c^n\right)^2, \tag{5.8}$$

where l_c^n denotes the c-th element of the label vector of the n-th sample and similarly y_c^n is the c-th element of the network's output vector for the n-th input.

Let define the output $\mathbf{o}$ of a layer ℓ as:

$$\mathbf{o}^\ell = f(\mathbf{u}^\ell), \mathbf{u}^\ell = \mathbf{W}^\ell \mathbf{o}^{\ell-1} + b^\ell, \tag{5.9}$$

where function $f(\cdot)$ denotes again a non-linear activation function. For fully connected layers, the back-propagated error δ^ℓ is computed as:

$$\delta^\ell = \left(\left(\mathbf{W}^{\ell+1}\right)^T \delta^{(\ell+1)}\right) \circ f'\left(\mathbf{u}^\ell\right), \tag{5.10}$$

where $\circ$ stands for the element-wise multiplication, and $f'(\cdot)$ is the derivative of the activation function. For the output layer L, the error is computed as:

$$\delta^L = f'\left(\mathbf{u}^\ell\right) \circ \left(\mathbf{y}^n - \mathbf{l}^n\right), \tag{5.11}$$

The gradients are computed as:

$$\nabla_{W^\ell} E = \mathbf{o}^{\ell-1} \left(\delta^\ell\right)^T, \tag{5.12}$$

$$\nabla_{b^\ell} E = \delta^\ell. \tag{5.13}$$

The weights and biases are updated according to the rules:

$$\Delta \mathbf{W}^\ell = -\eta \nabla_{W^\ell} E, \tag{5.14}$$

$$\Delta b^\ell = -\eta \nabla_{b^\ell} E, \tag{5.15}$$

respectively, where η expresses the learning rate. For paired convolutional and subsampling layers, the error of the j-th kernel in ℓ-th layer is estimated as:

$$\delta_j^\ell = up\left(\left(\mathbf{W}_j^{\ell+1}\right)^T \delta_j^{(\ell+1)}\right) \circ f'\left(\mathbf{u}_j^\ell\right). \tag{5.16}$$

where up is a simple function which upsamples the input by tiling each pixel in horizontal and vertical direction. Its upsampling factor is the same as the subsampling factor of the current pooling layer. It calculates the error with respect to the subsampled units. Gradients are computed as:

$$\nabla_{k_{ij}^\ell} E = \sum_{u,v} \left(\delta_j^\ell\right)_{uv} \left(p_i^{\ell-1}\right)_{uv}, \tag{5.17}$$

$$\nabla_{b_j} E = \sum_{u,v} \left(\delta_j^\ell\right)_{uv}, \tag{5.18}$$

where $p_i^{\ell-1}$ denotes a patch from $o_i^{\ell-1}$ that was multiplied element-wise by a kernel k_{ij}^ℓ during the convolution while the element at the position (u, v) of the output convolution map x_j^ℓ was computed.

Since convolutional kernels are applied over the whole input maps, the number of connections is much higher than the number of weights. This means that weighs are shared and there are fewer parameters than in ordinary NNs.

Proposed CNN architecture

The CNN architecture used in this work is depicted in Fig. 5.19. It consists of two convolutional and two mean-pooling layers in alternating order. In both convolutional layers, 24 convolutional filters of kernel size 5×5 are used. The input of the network is an image patch of size $4 \times d \times d$ (four MR sequences are present in multisequence volumes) and the output is a vector of length N indicating membership to one of the N classes in the label patch dictionary.
An example of patch propagation through the proposed network is depicted in Fig. 5.20.

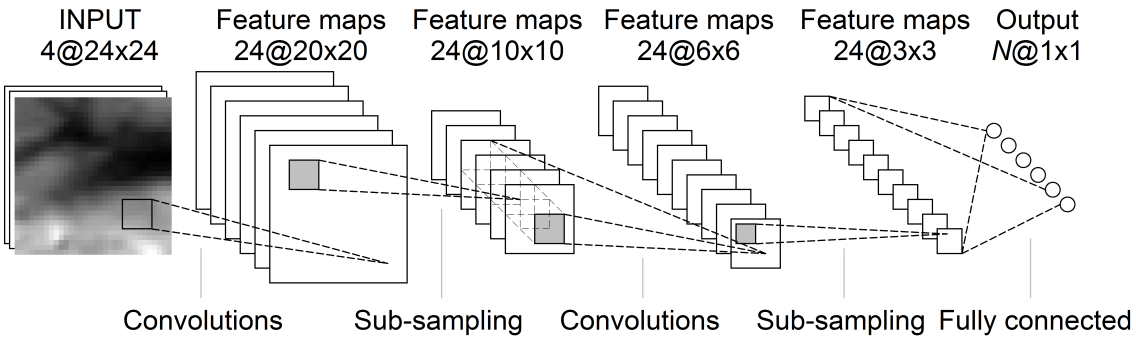


Fig. 5.19: Architecture of the CNN for $d = 24$. The input of the network is a multisequence image patch. The output of the network are N probabilities, where N denotes the size of a label patch dictionary.

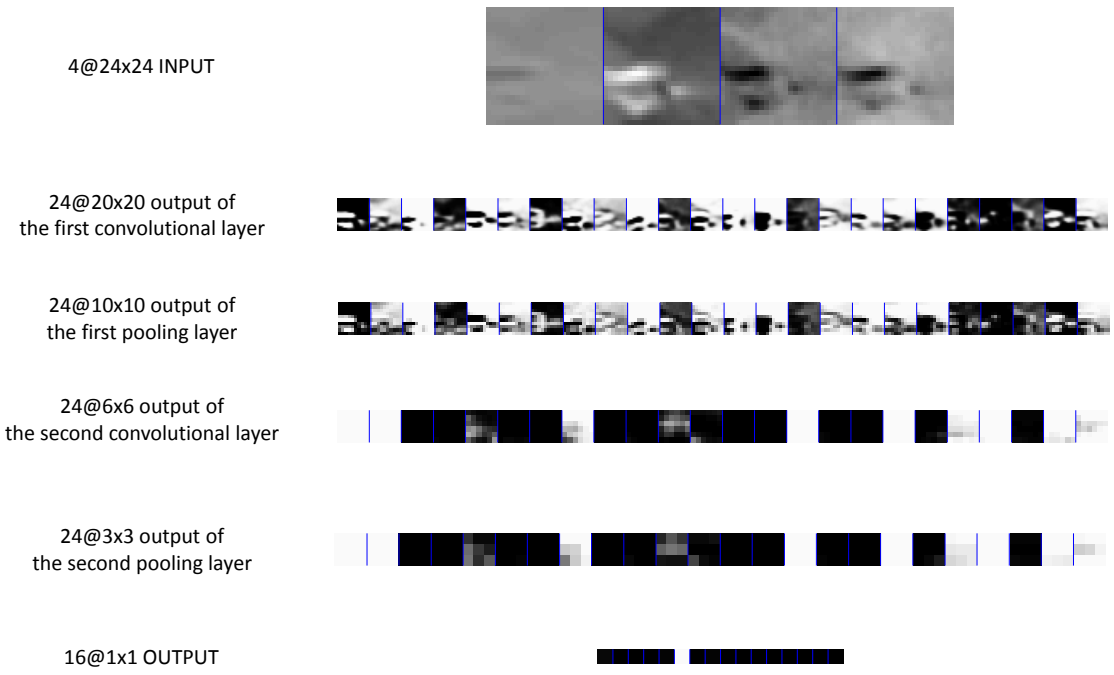


Fig. 5.20: Example of a patch propagation through the network and the outputs of each network layer.

5.4.3 Local structure prediction

This section describes a novel approach for classification-based medical image segmentation techniques, which lies in the principled combination of a deep learning approach together with local structure prediction in the medical image segmenta-

tion task. This approach takes the advantage of the fact that most medical images feature a high similarity in the intensities of nearby pixels and a strong correlation of intensity profiles across different image modalities. One way of dealing with – and even exploiting – this correlation is the use of local image patches. In the same way, there is a high correlation between nearby labels in image annotation, a feature that is used in the "local structure prediction" of local label patches.

Let $\mathbf{x}$ be the *image patch* of size $d \times d$ from the image space $\mathcal{I}$. Focusing on 2D patches, a patch $\mathbf{x}$ is represented as $\mathbf{x}(u, v, I)$ where (u, v) denotes the patch top left corner coordinates in a multisequence image $I(s, V)$ where s denotes the slice position in a multisequence volume V.

Label patches

Treating the annotation task for each class individually, a label space $\mathcal{L} = \{0, 1\}$, which is given by an expert's manual segmentation of the pathological structures, is obtained. The *label patch* is then a patch $\mathbf{p}$ of size $d' \times d'$ from the structured label space $\mathcal{P}$, i.e. $\mathcal{P} = \mathcal{L}^{d' \times d'}$. The label size d' is equal or smaller than the image patch size d. The label patch $\mathbf{p}$ is centered on its corresponding image patch $\mathbf{x}$ (Fig. 5.21), and it is represented as $\mathbf{p}(u+m, v+m, L)$ where $L(s, W)$ is a manual segmentation in a slice s of a label volume W and m denotes the margin defined as $m = \frac{1}{2}(d - d')$.

Optimal values for d and d' and, hence, the ratio $r = \frac{d'}{d}$ may vary depending on the structure to be segmented and the image resolution.

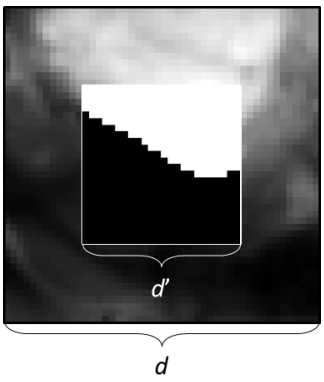
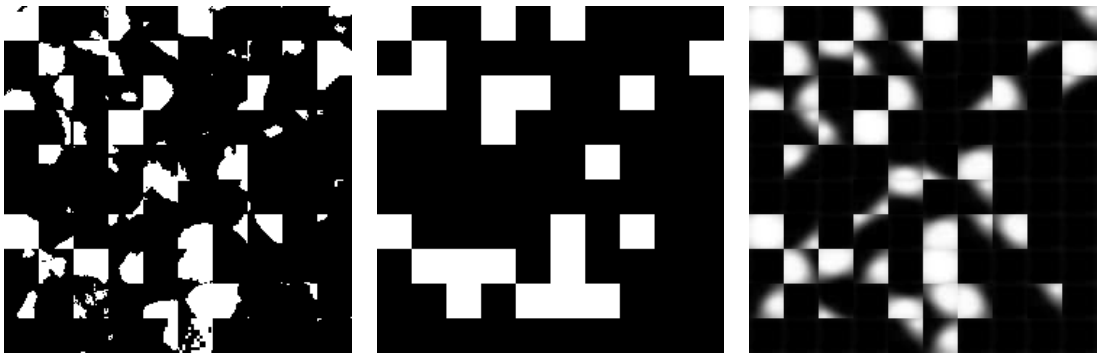


Fig. 5.21: Local structured prediction: Image feature patches (with side length d) are used to predict the most likely label patch (with side length d') in its center. While standard patch based prediction approaches use $d' = 1$ (voxel), the proposed approach considers all values with $1 \leq d' \leq d$.

Generating the label patch dictionary

Label patches $\mathbf{p}$ are clustered into N groups using k-means leading to a label patch dictionary of size N. Subsequently, the *label template* $\mathbf{t}$ of a group n is identified as the average label patch of a given cluster. In the segmentation process, these smooth label templates $\mathbf{t}$ are then used for the segmentation map computation rather than strict border prediction as used in previous local structure prediction methods [16][58][135]. The structures are learned directly from the training data instead of using predefined groups as in [135]. An example of ground truth label patches with their representation by a dictionary of size $N = 2$ (corresponding to common segmentation approach) and $N = 32$ is depicted in Fig. 5.22.



(a) (b) (c)

Fig. 5.22: Ground truth label patches (a) with corresponding binary (b) and structured (c) representation.

The size of the label patch dictionary N and, hence, the number of classes in the classification problem, may differ between problems depending on variability and shape complexity of the data. Figure 5.23 shows an example of clustering training labels patches into 16 classes. The left side of the figure depicts the average labels of each cluster, while the right side shows a subset of train labels assigned to a cluster with a given average patch, i.e. the labels that would be replaced by the same average label in the original image.

Defining the N-class prediction problem

After having obtained a set of N clusters, the binary segmentation problem is transformed into an N class prediction task: Each image patch $\mathbf{x}$ is identified in the training set with the group n that the corresponding label patch $\mathbf{p}$ has been assigned to during the label patch dictionary generation. In prediction, the label template $\mathbf{t}$ of the predicted group n (size $d' \times d'$) is assigned to the location of each

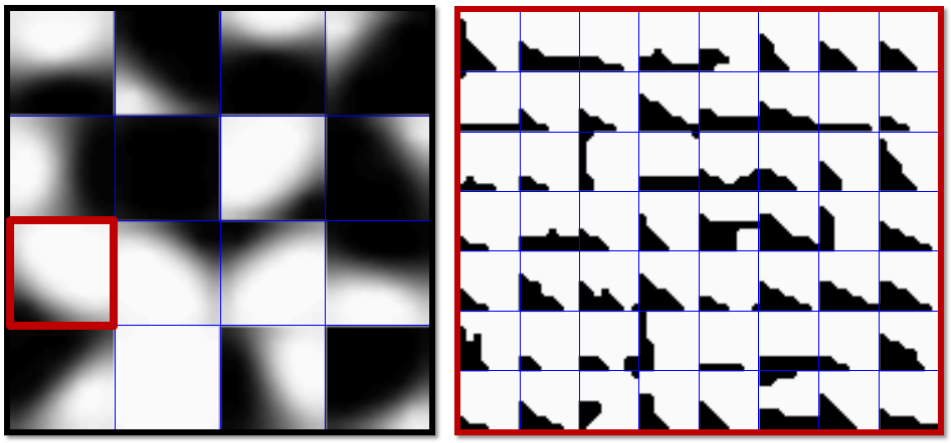


Fig. 5.23: Example of clustering the training labels patches into 16 groups. Left: average labels estimated for each group. Right: subset of 56 train labels assigned to a group with the average patch highlighted by a red bounding box in the left image.

image patch and all overlapping predictions of the neighborhood are averaged. According to the experiments, a discrete threshold $th = 0.5$ was chosen for final label prediction.

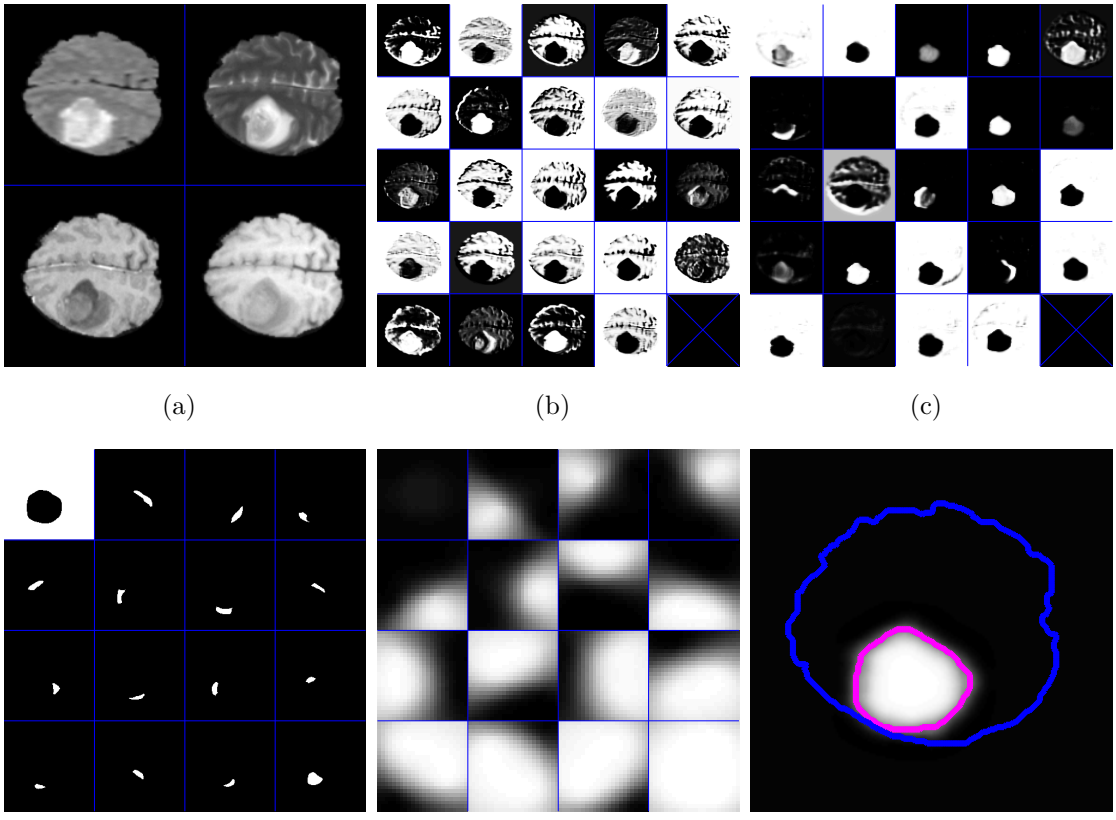
5.4.4 Slice Inference

Image patches from multisequence volumes are mapped into m 2D input channels of the network, where m denotes the number of sequences. This is similar to RGB image mapping. During the training phase, patches of a given size are extracted from training volumes. Using the same approach for testing is inefficient and therefore different approach used in [89] is employed instead. The whole input multisequence 2D slice is fed to the network architecture, which leads to a significantly faster convolutional process than applying the same convolution several times to small patches. This requires proper slice padding to be able to label pixels close to the slice border.

The output of the network is a map of label scores. However, this label map is smaller than the input slice due to the presence of pooling layers inside the CNN architecture. Two 2×2 pooling layers are present in the proposed architecture, which

means that there is only one value for every 4×4 region. Pinheiro and Collobert [89] fed the network by several versions of input image shifted on X and Y axis and merged the outputs properly. More common approach is to upscale the label map to the size of the input image. The latter approach is faster since only one convolution per slice is performed compared to 16 when using the former approach in case of the proposed CNN architecture. However, both of them were tested and compared.

One can see the sequential processing of the input multisequence slice in Fig. 5.24. 5.24(b) and 5.24(c) depict 24 outputs of the first and the second convolutional layer of the proposed CNN architecture. 5.24(d) shows the final classification map of the network. Note the average labels for each group in 5.24(e). One can compare them to the ground truth tumor border in the input image. The final probability map of the whole tumor area is depicted in 5.24(f).



(a) (b) (c)

(d) (e) (f)

Fig. 5.24: Sequential processing of multisequence slice (a). (b) and (c) show all 24 outputs of the first and the second convolutional layer of the proposed CNN architecture. (d) depicts the output of the whole network for 16 groups with the average patch labels depicted in (e). (f) shows the final probability map of the whole tumor area with outlined brain mask (blue) and final segmentation (magenta) obtained by thresholding by $th = 0.5$.

Since a hierarchy exists between particular segmentation sub-tasks, both tumor core and active tumor are segmented only inside the whole tumor region. This makes the segmentation process significantly faster. Although the hierarchy exists between tumor core and active tumor too, this approach is not used here since the segmentation of tumor core is the most difficult sub-task and it is usually the least accurate one.

6 RESULTS AND DISCUSSION

This chapter describes all tests that were performed for all three algorithms. The description of the experiments is in the same order as the algorithms description in Chap. 5, i.e. supervised 2D brain tumor presence detection (Sec. 6.3), unsupervised 2D and 3D brain tumor extraction (Sec. 6.4 and 6.5), and supervised brain tumor segmentation in 3D volumes using local structure prediction (Sec. 6.6). The evaluation criteria and experimental setup are described in Sec. 6.2. The extraction and segmentation algorithms use the same criteria, while only several of them are used to evaluate the presence detection algorithm. The overall algorithm performance as well as the performance for high-grade (HG) and low-grade (LG) gliomas is evaluated. For each test, the average computing time for the test database is also mentioned. These times do not include the inhomogeneity correction, skull stripping and image co-registration since all data sets had been preprocessed before they were released. All experiments were run on 4-core CPU Intel Xeon E3 3.30GHz without GPU acceleration using Matlab.

All algorithms were tested on publicly available data from the MICCAI 2014 Challenge on Multimodal Brain Tumor Image Segmentation (BRATS [1]. The description of the data is given in the first section of this chapter (Sec. 6.1).

6.1 Dataset

Brain tumor image data used in this work were obtained from the training data set of the MICCAI 2014 Challenge on Multimodal Brain Tumor Image Segmentation (BRATS) organized by Bjoern Menze, TU Munchen; Mauricio Reyes, Bern University; Keyvan Farahani, NIH; and Jayashree Kalpathy-Cramer, Harvard MGH. The challenge database contains fully anonymized images from the following institutions: ETH Zurich, University of Bern, University of Debrecen, and University of Utah and publicly available images from the NIH Cancer Imaging Archive (TCIA). For each patient, T1, T2, FLAIR, and post-Gadolinium T1 MR volumes are available. All volumes are linearly co-registered to the T1C image, skull stripped, and interpolated to 1mm isotropic resolution. No attempt was made to put the individual patients in common reference space.

Annotations comprise the whole tumor, the tumor core (including cystic areas), and the Gd-enhanced tumor core and are described in the BRATS reference paper [80] recently published in IEEE Transactions for Medical Imaging.

[1] http://www.braintumorsegmentation.org

All data sets used in this work originating from the BRATS2012 and BRATS2013 challenge were segmented manually (by four raters). Training data from TCIA were annotated by fusing results of segmentation algorithms that ranked high in the BRATS 2012 and BRATS 2013 challenge. Annotations were inspected visually and were approved by experienced raters.

An example of a manually annotated volume is given in Fig. 6.1. The figure shows manual annotations for three tumor structures (whole tumor, tumor core, and active tumor) in different MR sequences and their fusion. Note the hierarchy (active tumor $\subseteq$ tumor core $\subseteq$ whole tumor) between the structures.

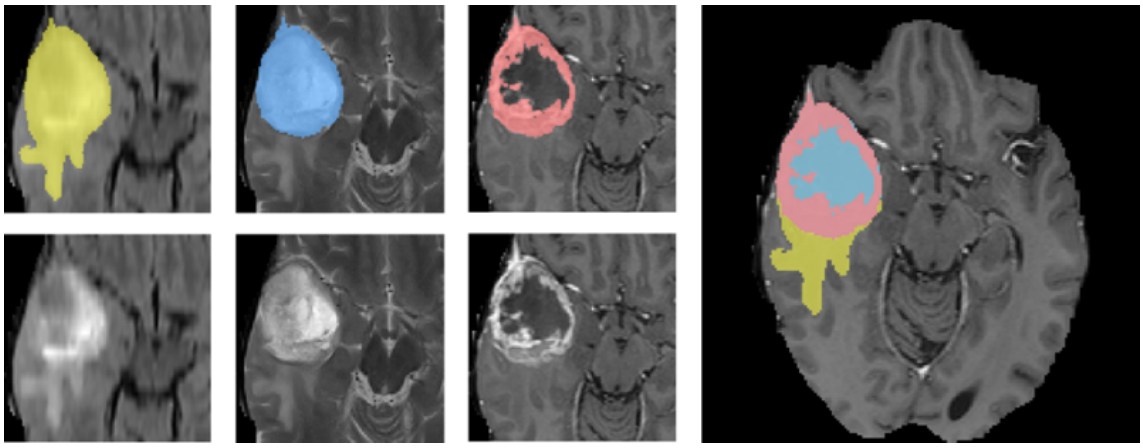


Fig. 6.1: Example of a manually annotated volume. Left images show a cut-out with brain tumor of the same area in three different MR sequences with three different annotated structures (from left to right: whole tumor, tumor core, and active tumor). The most right image shows the fusion of the annotations in the whole slice.

The data set used for the evaluation of the proposed algorithm contains 254 real multisequence volumes of 199 high-grade and 55 low-grade glioma subjects.

For brain tumor presence detection and brain tumor segmentation using local structure prediction, the data set was divided into three groups: training, validation and testing. The training set consists of 130 high-grade and 33 low-grade glioma subjects, the validation set consists of 18 high-grade and 7 low-grade glioma subjects, and the testing set consists of 51 high-grade and 15 low-grade glioma subjects, summing up to 254 multisequence volumes of average size $240\times240\times155$. For unsupervised brain tumor extraction, all 254 volumes were used for testing since no training phase is required there.

6.2 Evaluation criteria

Segmentation algorithms are usually evaluated by the Dice score (DS). Dice score, sometimes called Dice similarity coefficient, is computed according to the equation:

$$DS = \frac{2\,|A \bigcap B|}{|A| + |B|}, \tag{6.1}$$

where A and B denote the ground truth and the extraction result masks, respectively. The range of the DS values is [0;1], where DS = 1 expresses the exact similarity.

Other measures widely employed for segmentation accuracy are Precision, which is sometimes called Positive predictive value (PPV), and Sensitivity, sometimes called Recall. They are both based on the Confusion matrix, a 2×2 (in case of binary classification) matrix:

		Truth	
		True	False
Prediction	True	TP	FP
	False	FN	TN

where TP, FP, FN and TN stand for "True Positive", "False Positive", "False Negative", and "True Negative", respectively.

The presence detection performance is evaluated by the confusion matrix, which, for easier interpretation, is modified to:

		Truth	
		True	False
Prediction	True	TPR	FPR
	False	FNR	TNR

where TPR, FPR, FNR and TNR stand for "True Positive Rate" defined as $\frac{TP}{TP+FN}$, "False Positive Rate" defined as $\frac{FP}{FP+TN}$, "False Negative Rate" defined as $\frac{FN}{TP+FN}$, and "True Negative Rate" defined as $\frac{TN}{FP+TN}$, respectively.

Precision is computed as:

$$Precision = \frac{TP}{TP + FP}, \tag{6.2}$$

which measures the percentage of the true positives in the scope of all positive predictions. Sensitivity, on the other hand, measures the percentage of the true positives in the scope of all real positives and is defined as:

$$Sensitivity = \frac{TP}{TP + FN}, \tag{6.3}$$

which corresponds to the TPR. Accuracy, which is defined as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \tag{6.4}$$

will only be used for the tumor presence detection algorithm. This measure is not suitable for brain tumor segmentation evaluation, which is a very imbalanced task because the number of negatives is much higher than the number of positives. Accuracy always reaches high value here.

6.3 Brain tumor presence detection

In this section, the method proposed for detection of particular structures of brain tumor, i.e. whole tumor, tumor core, and active tumor, is evaluated. This method is based on a novel approach using multiresolution symmetry analysis described in Sec. 5.2.

6.3.1 Experimental setup

This algorithm was tested for all four MR sequences, i.e. FLAIR, T2, T1 and T1C. Since brain symmetry exists in two planes, axial and coronal, both of them were considered and tested separately during the algorithm evaluation. Note that slices with tumor core and active tumor are also included in the set of slices with whole tumor; therefore the total number of slices is equal to sum of whole tumor and healthy slices.

The numbers of pathological slices (with a pathology of size $\geq$ 100px, i.e. a pathology of size $\geq$ 1cm^2 considering the 1mm isotropic resolution) and healthy slices (with no tumor region) in axial and coronal planes are summarized in Tab. 6.1 and Tab. 6.2, respectively. This table shows the distribution of the data set in training, validation and testing subsets. The distribution follows the division described in Sec. 6.1 but it shows the exact numbers of slices used during the algorithm evaluation.

The total number of extracted features for stand-alone images was 24 per sequence, i.e. a maximum from multiresolution asymmetry map and three maxima from each single-resolution map, five absolute thresholding levels, and five relative

Tab. 6.1: Experimental setup of brain tumor presence detection in axial plane.

	Whole tumor	Tumor core	Active tumor	Healthy	Overall
Train	10572	6741	4910	9820	20392
Validation	1553	1057	708	1639	3192
Test	4335	3099	2442	4268	8603
Overall	16460	10897	8060	15727	32187

Tab. 6.2: Experimental setup of brain tumor presence detection in coronal plane.

	Whole tumor	Tumor core	Active tumor	Healthy	Overall
Train	9887	4369	1922	11323	21210
Validation	1449	758	307	1883	3332
Test	4176	2353	1079	4859	9035
Overall	15512	7480	3308	18065	33577

thresholding levels. Each thresholding leads to two features: number of regions and size of all regions. In 3D application, four more intensity features were computed for each threshold: average intensity in created regions, absolute value of the difference between average intensity in regions in the left and the right hemisphere, standard deviation of intensities in created regions, and absolute value of the difference between intensity standard deviation in regions in the left and the right hemisphere. This sums up to 64 features for slices in 3D volume.

6.3.2 Application to the test set

Stand-alone slices

Axial plane. The accuracy of whole tumor, tumor core, and active tumor detection in axial slices is shown in Tab. 6.3. In this table, one can see that in axial plane, the best accuracy for the whole tumor region is reached in FLAIR images, for the tumor core and active tumor regions in FLAIR or T2 images. When all MR sequences are combined, the accuracy is improved for all detected structures.

Confusion matrices for multisequence detection of a given part of the brain tumor are shown in Tab. 6.4, 6.5 and 6.6.

Tab. 6.3: Accuracy of particular brain tumor structures detection in axial slices of different MR sequences.

Accuracy (in %)	Whole tumor	Tumor core	Active tumor
FLAIR	90	89	83
T2	88	89	83
T1C	83	84	81
T1	81	80	78
Multisequence	90	90	88

Tab. 6.4: Confusion matrix of active tumor presence detection in multisequence stand-alone axial slices.

	Tumor present	Tumor absent
Test positive	91	11
Test negative	09	89

Tab. 6.5: Confusion matrix of tumor core presence detection in multisequence stand-alone axial slices.

	Tumor present	Tumor absent
Test positive	88	08
Test negative	12	92

Tab. 6.6: Confusion matrix of active tumor presence detection in multisequence stand-alone axial slices.

	Tumor present	Tumor absent
Test positive	83	10
Test negative	17	90

Coronal plane. The accuracy of whole tumor, tumor core, and active tumor detection in coronal slices is shown in Tab. 6.7. In this table, one can see that in axial plane, the best accuracy for the whole tumor and active tumor regions is

reached in FLAIR images, for the tumor core region in FLAIR images. When all sequences are combined, the accuracy is improved for all detected structures.

Tab. 6.7: Accuracy of particular brain tumor structures detection in coronal slices of different MR sequences.

Accuracy (in %)	Whole tumor	Tumor core	Active tumor
FLAIR	90	88	86
T2	85	89	86
T1C	80	83	81
T1	76	77	75
multisequence	91	91	88

Confusion matrices for the multisequence detection of a given part of the brain tumor are shown in Tab. 6.8, 6.9 and 6.10.

Tab. 6.8: Confusion matrix of whole tumor presence detection in multisequence stand-alone coronal slices.

(in %)	Tumor present	Tumor absent
Test positive	91	09
Test negative	09	91

Tab. 6.9: Confusion matrix of tumor core presence detection in multisequence stand-alone coronal slices.

(in %)	Tumor present	Tumor absent
Test positive	87	07
Test negative	13	93

Presence detection in 3D volume

Axial plane. This test used the same experimental setup as the evaluation of presence detection in stand-alone images. However, image features are extracted together with the asymmetry features used in stand-alone image, because intensities

Tab. 6.10: Confusion matrix of active tumor presence detection in multisequence stand-alone coronal slices.

(in %)	Tumor present	Tumor absent
Test positive	81	09
Test negative	19	91

can be normalized in whole 3D volumes. The accuracy of whole tumor, tumor core, and active tumor detection in axial slices are shown in Tab. 6.11. One can see that in axial plane, the best accuracy for the whole tumor regions as well as for the tumor core region is reached in FLAIR images. Active tumor is most accurately detected in T1C slices.

Tab. 6.11: Accuracy of particular brain tumor structures detection in axial slices of different MR sequences.

Accuracy (in %)	Whole tumor	Tumor core	Active tumor
FLAIR	94	92	88
T2	89	90	88
T1C	83	87	89
T1	80	80	78
Multisequence	93	93	92

Confusion matrices for multisequence detection of a given part of the brain tumor are shown in Tab. 6.12, 6.13 and 6.14.

Tab. 6.12: Confusion matrix of active tumor presence detection in multisequence axial slices in 3D volume.

(in %)	Tumor present	Tumor absent
Test positive	95	09
Test negative	05	91

Coronal plane. The accuracy of whole tumor, tumor core, and active tumor detection in coronal slices are shown in Tab. 6.15. In this table, one can see that in

Tab. 6.13: Confusion matrix of tumor core presence detection in multisequence axial slices in 3D volume.

(in %)	Tumor present	Tumor absent
Test positive	92	06
Test negative	08	94

Tab. 6.14: Confusion matrix of active tumor presence detection in multisequence axial slices in 3D volume.

(in %)	Tumor present	Tumor absent
Test positive	92	08
Test negative	08	92

axial plane, the best accuracy for the whole tumor region as well as for the tumor core region is reached in FLAIR images. Active tumor is most accurately detected in T1C slices.

Tab. 6.15: Accuracy of particular brain tumor structures detection in coronal slices of different MR sequences.

Accuracy (in %)	Whole tumor	Tumor core	Active tumor
FLAIR	93	91	87
T2	86	90	87
T1C	80	87	88
T1	77	80	77
Multisequence	94	93	92

Confusion matrices for the multisequence detection of a given part of the brain tumor are shown in Tab. 6.16, 6.17 and 6.18.

Discussion

In this part, the supervised method for brain tumor presence detection was tested. It was evaluated for two different cases, the detection in stand-alone 2D images and slice-wise detection in 3D volumes. The former used only features extracted from

Tab. 6.16: Confusion matrix of whole tumor presence detection in multisequence coronal slices in 3D volume.

	Tumor present	Tumor absent
Test positive	90	06
Test negative	10	94

Tab. 6.17: Confusion matrix of tumor core presence detection in multisequence coronal slices in 3D volume.

	Tumor present	Tumor absent
Test positive	91	06
Test negative	09	94

Tab. 6.18: Confusion matrix of active tumor presence detection in multisequence coronal slices in 3D volume.

	Tumor present	Tumor absent
Test positive	95	09
Test negative	05	91

the asymmetry maps, while the latter added the asymmetry-based image intensity features. Figure 6.2 shows Precision-recall curves for axial plane test set for each tumor part detection using different MR sequences in both stand-alone slices and slices in 3D volume. Recall is equivalent to sensitivity. Green dashed curves depict isolevel lines of F-measure, which is equivalent to the Dice score. Graphs for the coronal plane are omitted since they are very similar to the graphs for the axial plane. As it can be concluded from the results, the intensity features bring important information, which in most tests improved the detection accuracy, e.g. the performance increased from 90% to 94% using axial FLAIR slices for whole tumor detection, or from 81% to 88% using coronal T1C slices for active tumor detection.

According to the results, it can be stated that in stand-alone images all parts can be automatically detected with the highest accuracy in FLAIR. In 3D volume slices, incorporating intensity-based features improves the results and active tumor is than detected with highest accuracy in T1C slices. However, using all MR sequences

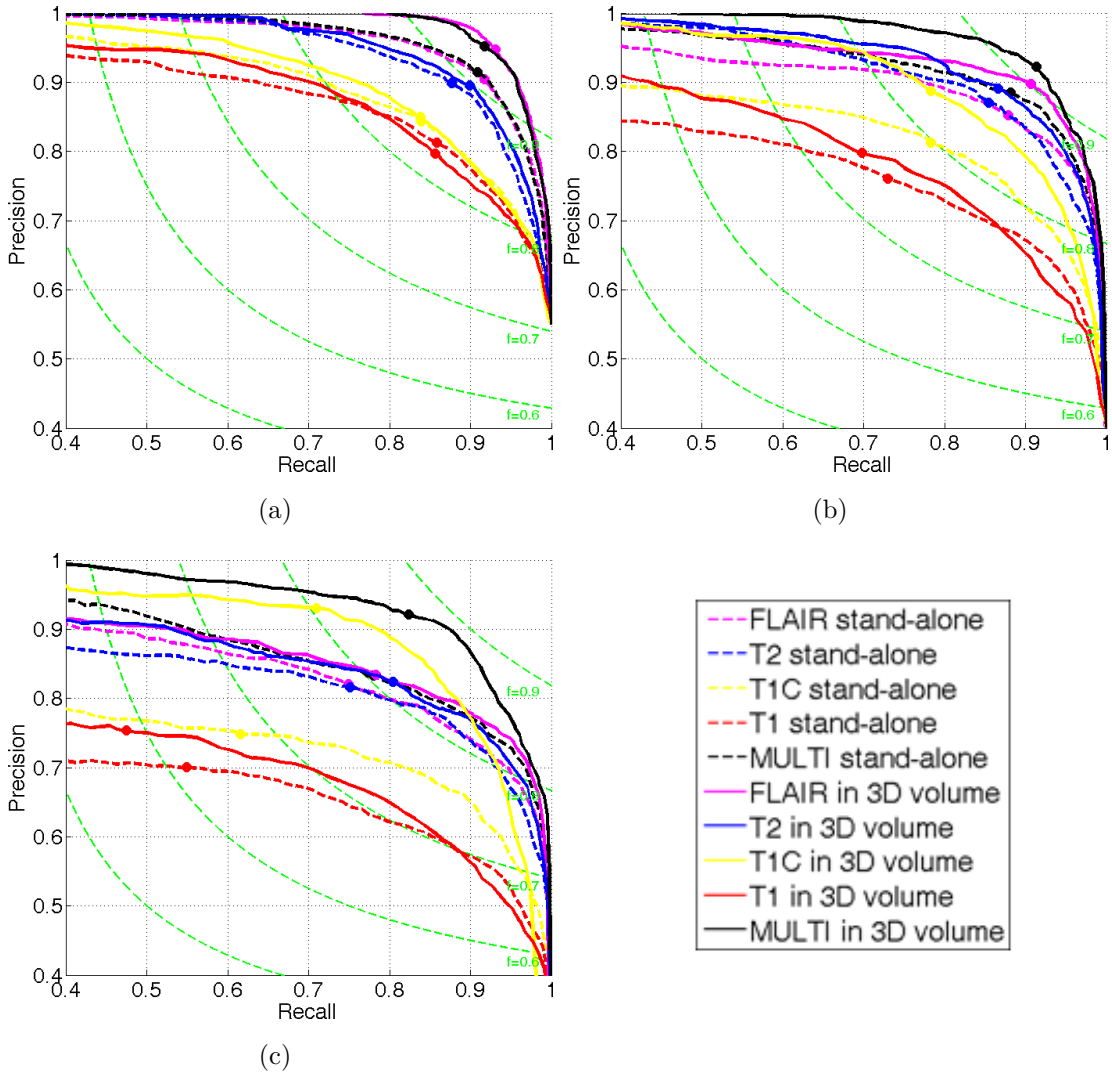


Fig. 6.2: Precision-recall curves for (a) whole tumor, (b) tumor core and (c) active tumor presence detection in both stand-alone slices (dashed curves) and slices in 3D volume (solid curves) for the axial plane test set. Circles on each curve represent the actual achieved results. Green dashed curves depict isolevel lines of F-measure, which is equivalent to the Dice score.

improves the performance in almost all detection tasks. It was shown that the algorithm reaches high accuracy in both axial and coronal planes. Since, according to the best knowledge of the author, there are no methods for the fast brain tumor presence detection task, it is not possible to compare the results with state-of-the-art algorithms.

The computing time of 0.11s per slice and sequence, where 90% of the time is consumed by symmetry analysis and the rest by feature extraction and classification, shows that this method can be used either for fast decision, whether an image

contains a tumor, or in the segmentation process as a pre-processing step to determine the slices to be segment or analyzed. Many state-of-the-art methods based on deep learning architectures use slice-wise approach, and, therefore, this method can also be used in pipelines of such algorithms.

6.4 Unsupervised 2D brain tumor extraction

After the whole tumor presence is detected in 2D, it can be extracted using the novel unsupervised 2D brain tumor extraction method described in Sec. 5.3. This method is designed for FLAIR images only, therefore, other MR sequences will not be considered here.

6.4.1 Experimental setup

Since the proposed algorithm is unsupervised, i.e. no training phase is required; it was tested on the whole set of 254 FLAIR volumes. Axial and coronal slices containing a whole tumor area of at least 6cm^2 were extracted from each volume since the method cannot deal with smaller pathologies. The overall numbers of axial and coronal slices included in the test are 22213 (18445 of HG and 3768 of LG) and 28078 (23401 of HG and 4669 of LG), respectively.

6.4.2 Application to the test set

The results of the algorithm application to axial and coronal slices are summarized in Tab. 6.19. The table contains the overall results as well as the results achieved for high- and low-grade gliomas evaluated by the Dice score, precision and sensitivity.

In the table, one can see higher sensitivity values than precision. This points to the fact that the resulting area is usually larger than the true tumor area, in other words, more false positives than false negatives are present. The sensitivity values show that in average 87% of positive pixels were found, while the precision values show that in average 64% of positive labels were truly positive. The average computing time per slice is 0.13 seconds, where 75% of the time is consumed by the symmetry analysis. Examples of the comparison of a manual annotation and an automatic extraction in both axial and coronal planes are shown in Fig. 6.3. Three successful and one unsuccessful (the most right) automatic segmentations are shown for each plane. The unsuccessful segmentation in axial plane shows the failure of the method in case of a small pathology. The algorithm was not able to correctly determine the threshold due to the small amount of tumorous pixels. The unsuccessful segmentation in coronal plane shows a failure of the method for

Tab. 6.19: Segmentation results reporting average and median Dice score, precision and sensitivity. Shown are results for axial and coronal slices. "std" and "mad" denote standard deviation and median absolute deviance. HG and LG stand for high- and low-grade gliomas, respectively.

	Axial		Coronal	
		HG / LG		HG / LG
Dice Score (in %)				
mean ± std	63±30	63±30 / 59±32	63±29	63±29 / 60±30
median ± mad	77±27	77±27 / 76±30	76±26	78±26 / 70±28
Precision (in %)				
mean ± std	64±37	64±37 / 57±39	63±37	65±37 / 58±38
median ± mad	85±35	85±35 / 71±37	81±34	84±34 / 65±36
Sensitivity (in %)				
mean ± std	87±31	87±16 / 91±12	87±15	86±15 / 91±12
median ± mad	92±11	91±12 / 96±09	91±11	91±12 / 95±09

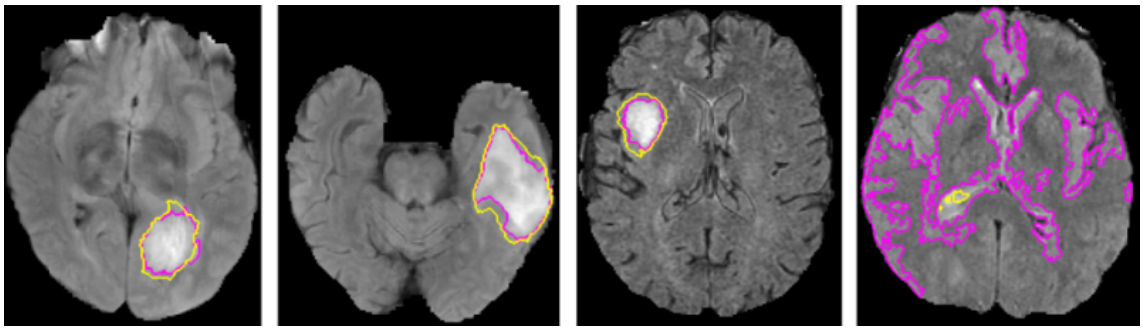
an image where the intensities of the pathology are very similar to the intensities of healthy tissues.

Discussion

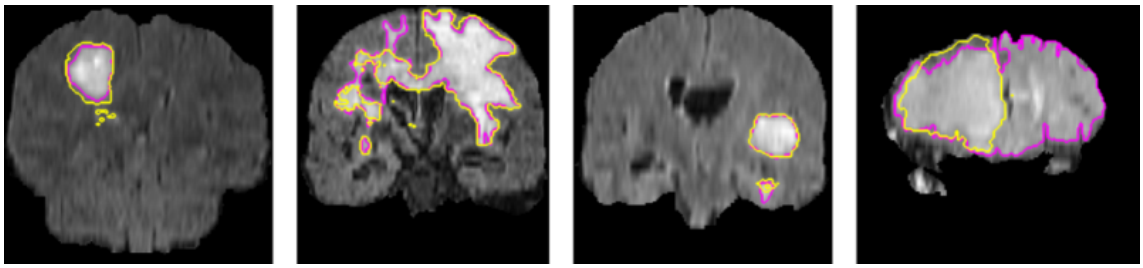
The purpose of this algorithm was to extract a tumor area (including edema) from a stand-alone FLAIR image. The results are not comparable to the state-of-the-art methods used for multisequence 3D brain tumor extraction, which are mostly based on machine learning techniques using the intensity information. Such approach is possible for 3D volumes where the intensity normalization can be performed. However, this is not suitable for a stand-alone axial or coronal slice with unknown information about the tumor size and the slice coordinates. This method was designed to deal with such images and reached promising results with average and median Dice scores of 63±30 and 77±27 and computing time of only 0.13 seconds.

6.5 Unsupervised 3D brain tumor extraction

In this section, the method proposed for unsupervised brain tumor locating and extracting is evaluated. This method is based on a novel unsupervised algorithm combining 3D multiresolution symmetry analysis described in Sec. 5.2 and automatic thresholding.



(a)



(b)

Fig. 6.3: Examples of the brain tumor extraction using the unsupervised algorithm in 2D (a) axial and (b) coronal FLAIR slices. A comparison of manual annotation (yellow) and automatic extraction (magenta) is depicted.

6.5.1 Experimental setup

The 3D unsupervised brain tumor extraction algorithm was also tested on the whole database of 254 multisequence volumes since it does not require any training phase, and, hence, data partitioning is not necessary. This algorithm requires only T2 and FLAIR images, therefore other MR sequences are not considered in these tests. During testing, the accuracy of the automatic tumor locating using the symmetry analysis is evaluated together with the automatic whole tumor extraction. The tumor location is determined by the maximum of the multiresolution asymmetry map. This point shows the highest probability of the tumor location and, hence, it is able to detect slices in all three planes where there is the maximum probability of the tumor presence. If this point lies within the area of a tumor, it is considered as a correct locating. Since the asymmetry map is symmetrical according to the mid-sagittal plane, i.e. two maxima exist, either of them is considered. The second test evaluates the accuracy of the whole tumor region extraction. This is evaluated by the Dice score, precision and sensitivity.

6.5.2 Application to the test set

Brain tumor locating

This test was performed separately for each MR sequence, i.e. T2 and FLAIR. Results for both MR sequences are summarized in Tab. 6.20. The average computing time for one sequence is 6.7±0.8 seconds per volume.

Tab. 6.20: Brain tumor locating using asymmetry detection. The table expresses the percentage of correctly located brain tumors.

Accuracy (in %)	Overall	HG / LG
T2	90	90 / 90
FLAIR	95	96 / 95
T2+FLAIR	95	95 / 94

As one can see in the table, the proposed algorithm can automatically locate a tumor and show slices with a tumor within less than 7 seconds with an accuracy of 95% in FLAIR volumes and 90% in T2 volume of 1mm isotropic resolution. Combination of T2 and FLAIR volumes, where the time necessary for the asymmetry detection is twice higher, did not bring any improvement, and, hence, it can be stated that FLAIR volume is more suitable for brain tumor locating, at least in case of using the symmetry prior and the proposed approach.

An example of the tumor locating algorithm is shown in Fig. 6.4 where slices with the maximum asymmetry of all three planes are shown. The coordinates of this maximum are depicted by blue circle. The figure shows four different slices per subject. Two slices are shown in sagittal plane. The purpose is to show both slices with the maximum asymmetry, i.e. one slice from the left an one slice from the right hemisphere.

Brain tumor extraction

The performance of the brain tumor extraction algorithm based on the symmetry analysis is evaluated here by mean and median values of Dice score, precision and sensitivity. All results are summarized in Tab. 6.21. The overall performance as well as the performance for high-grade (HG) and low-grade (LG) gliomas is shown. The average computing time is 13.8±1.3 seconds per volume including the previous symmetry analysis in both volumes, which reaches about 93% of the computing time. Slightly lower performance was achieved by applying symmetry analysis only

to one volume, e.g. FLAIR. However, the computing time decreased to 7.7±0.9 seconds in this case.

Tab. 6.21: Brain tumor extraction using asymmetry detection.

	Overall	HG / LG
Dice Score (in %)		
mean ± std	66±21	65±22 / 70±20
median ± mad	74±17	74±17 / 75±15
Precision (in %)		
mean ± std	89±22	90±22 / 88±23
median ± mad	98±14	99±14 / 98±15
Sensitivity (in %)		
mean ± std	59±21	57±21 / 65±22
median ± mad	64±17	62±17 / 68±17

Examples of the comparison between the automatic extraction and the manual annotation are given in Fig. 6.4 where manual annotations are depicted by yellow color, and the results of the automatic extraction process are shown by magenta color. The figure shows four different slices per subject. Two sagittal slices are shown. The purpose is to show both slices with the maximum asymmetry, i.e. one slice from the left and one slice from the right hemisphere.

Discussion

In this section, the unsupervised algorithm for fast brain tumor locating and extraction was tested. The unsupervised approach for the brain tumor locating and delineating was verified on FLAIR and T2 volumes of 254 subjects. The proposed method used an approach different from other state-of-the-art algorithms. Compared to them, it is not supervised and, therefore, it does not require any training phase. It cannot reach the top three state-of-the-art results reported in [80] with 79%-82% (here: 64%), but it still reached comparative results with other methods described in that paper. However, the advantage of the proposed algorithm is its speed since it is not that demanding as other methods. With 13.8 seconds, it is faster than the fastest method reported in [80], where only 3 methods out of 16 running on CPU were able to extract brain tumor in less than 8 minutes. Therefore, the proposed algorithm is suitable for fast preliminary approximate tumor extraction rather than accurate segmentation. The first part of the algorithm, i.e. brain tumor

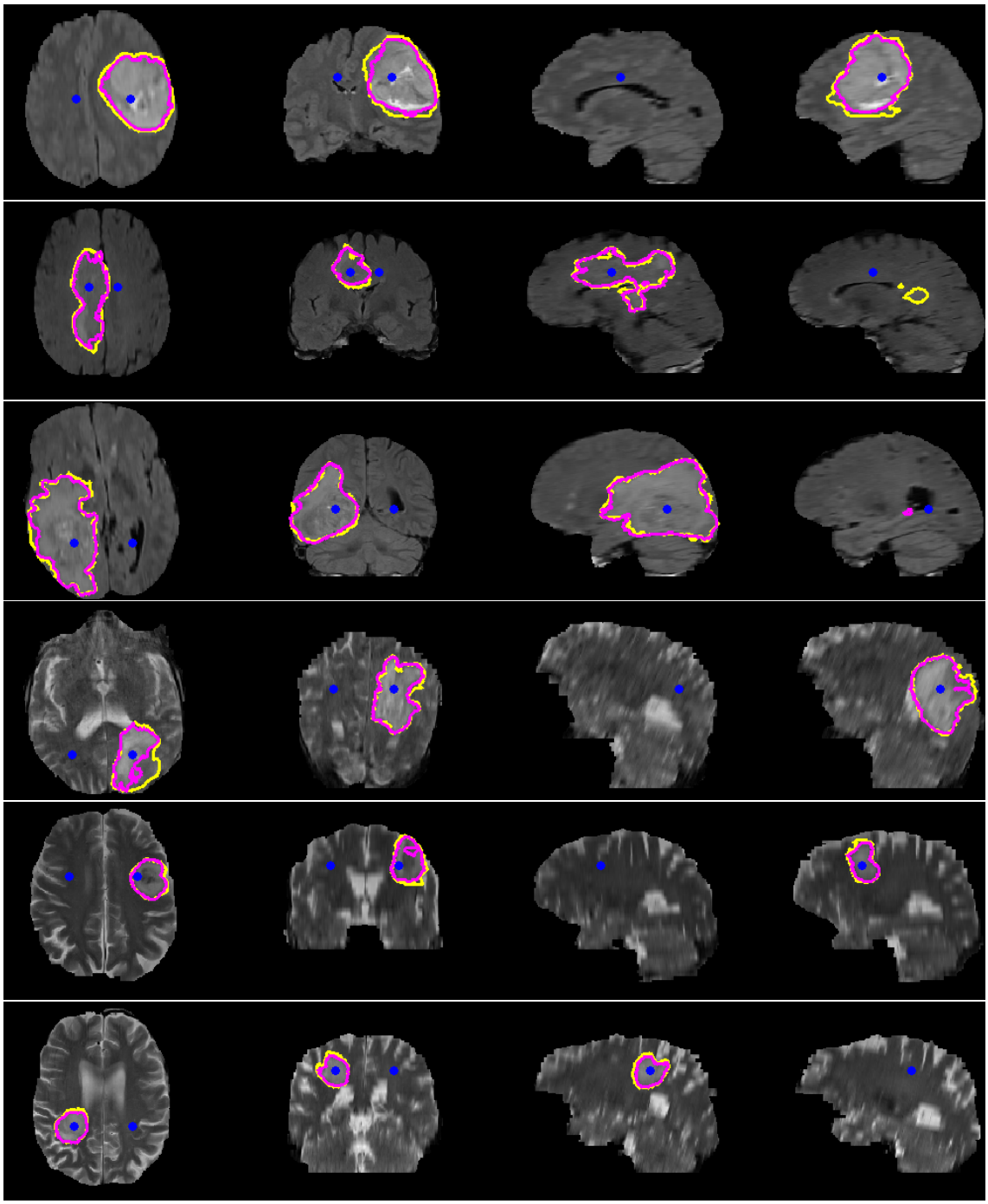


Fig. 6.4: Examples of the unsupervised brain tumor extraction based on the multiresolution symmetry analysis. The results are demonstrated on three FLAIR (upper rows) and three T2 (lower rows) volumes. The images show slices where the maximum asymmetry was detected with corresponding manual annotation (yellow) and automatic extraction (magenta). Blue circles point to the voxel with the highest asymmetry. Both sagittal slices with the maximum asymmetry are shown.

locating, is suitable for fast brain tumor locating for pre-analysis of a single-sequence volume since it may pre-analyze a single volume in less than 8 seconds.

6.6 Brain tumor segmentation using structure prediction

In this section, the method proposed for segmentation of particular structures of the brain tumor, i.e. whole tumor, tumor core, and active tumor, is evaluated. This method is based on an approach, whose novelty lies in the principled combination of the deep approach together with the local structure prediction in medical image segmentation task. The algorithm was described in Sec. 5.4.

6.6.1 Experimental setup

The algorithm is designed for a binary segmentation problem and it was applied separately to three brain tumor segmentation sub-problems: segmentation of whole tumor, tumor core and active tumor.

As it has been shown in [93], the computational demands of 3D CNN are still out of scope for today's computers. Therefore the volume is processed sequentially in 2D in the plane with the highest resolution, the axial plane here. Image patches from each multisequence volume are mapped into four 2D input channels of the network. This approach gives a good opportunity for parallelization of this task to reduce the computing time.

Alternatives to this basic approach have been proposed: slice-wise 3D segmentation approaches using CNN were proposed used in [93] and [101]. The former showed non-feasibility of using 3D CNN for larger cubic patches and proposed using of a 2D CNN for each orthogonal plane separately. The later proposed extraction of corresponding patches for a given pixel from each orthogonal plane and mapping them as separated feature maps. In this work, both of these approaches were tested and compared to the single slice approach that was chosen here.

From each training subject, 1500 random 2D multisequence image patches with corresponding label patches were extracted summing up to 244 500 training image patches. To ensure approximate balance of the database, higher probability of patch extraction was around the pathological area.

The parameters of the algorithm, i.e. the image patch size, the label patch dictionary size, and the label patch size were optimized separately for each subtask on the validation set and the results of the optimization process are described

in 6.6.2. It is followed by the description of the application to the tests set with examples of the segmentation results that are compared to manual annotations.

6.6.2 Parameter Optimization

Besides the parameters of the convolutional architecture, there are parameters of the proposed model: the image patch size d, the label patch size d', and the size of the label patch dictionary N. These parameters were tested with a pre-optimized fixed network architecture depicted in Fig. 5.19, which consisted of two convolutional layers, both with 24 convolutional filters of kernel size 5×5, and two mean-pooling layers in alternating order. The values selected for subsequent experiments are highlighted in graphs with a red vertical line.

Image patch size

The image patch size d is an important parameter since the segmented structures have different sizes and therefore less or more information is necessary for the label structure prediction. Figure 6.6.2 shows the Dice score values for different patch sizes with their best label patch size. According to the graphs, $d = 8$ was selected for the active part segmentation and $d = 24$ for the segmentation of tumor core and whole tumor. All three tests were performed for $N = 32$, which according to the previous tests is sufficiently enough for all patch sizes. The best results were in all cases achieved for $d' \geq \frac{1}{2}d$. The values selected for subsequent experiments are indicated by a red vertical line.

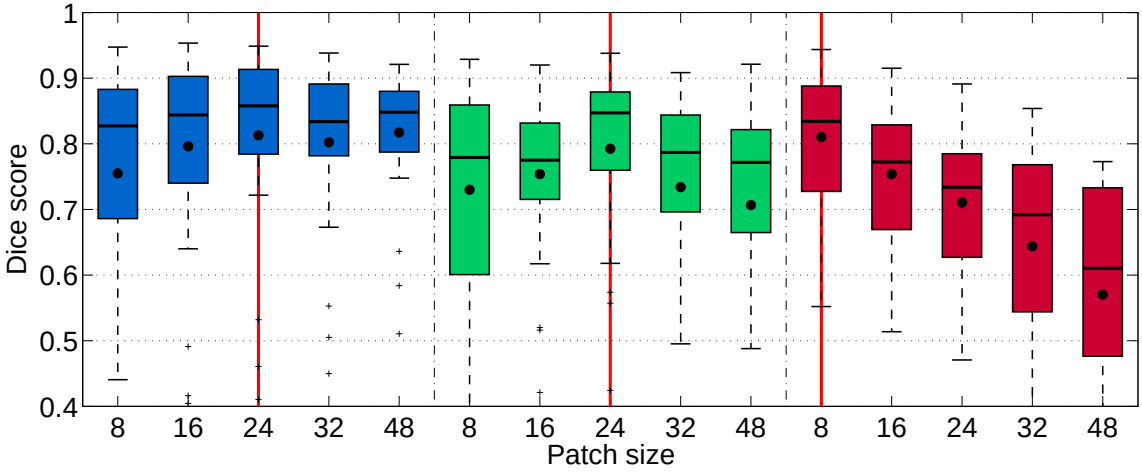


Fig. 6.5: Dice score as a function of the image patch size d with its best label patch size d' and the label patch dictionary size $N = 32$ for whole tumor (blue), tumor core (green) and active tumor (red).

Size of the label patch dictionary

The size of the label patch dictionary N influences differences between each label template $\mathbf{t}$ as well as the differences between belonging image patches $\mathbf{x}$ in each groups n. Results for several values of N are depicted in Fig. 6.6.2. Generally the best results were achieved for $N = 16$. The results were evaluated in similar manner as in the previous test, i.e. the best d' is used for each value of N. The values selected for subsequent experiments are indicated by a red vertical line.

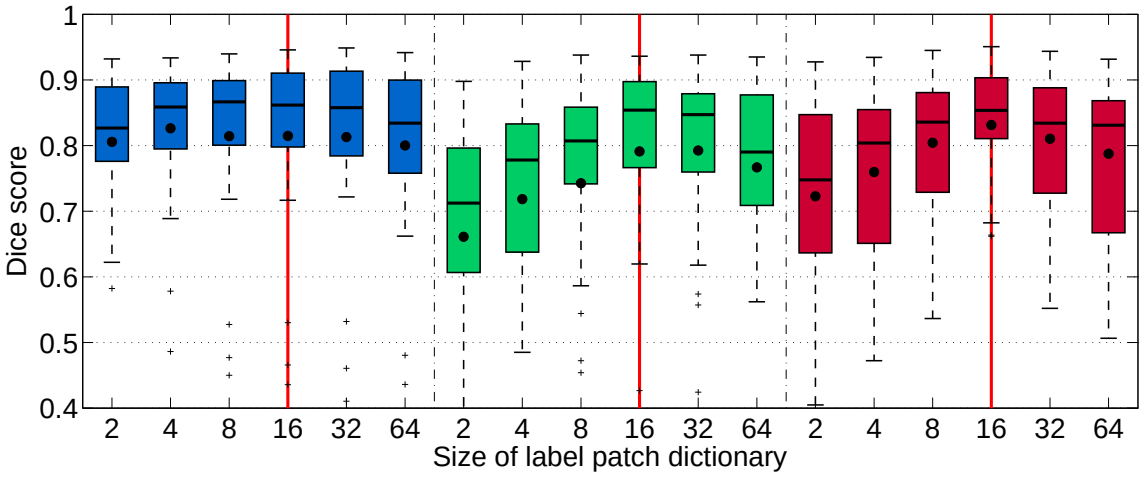


Fig. 6.6: Dice score as a function of the label patch dictionary size N using the optima of Fig. 6.6.2: $d = 24$ for whole tumor (blue), $d = 24$ for tumor core (green), $d = 8$ for active tumor (red).

Label patch size

The label patch size d' influences the size of structure prediction as well as the number of predictions for each voxel. Figure 6.6.2 shows the increasing performance with increasing d'. The values selected for subsequent experiments are indicated by a red vertical line.

2D versus 3D

Both triplanar and 2.5D deep learning approaches for 3D data segmentation as proposed in [93] and [101], respectively, were tested and compared to the single slice-wise segmentation approach. It was discovered that both approaches even decreased the performance of the proposed method by 2% and 5%, respectively.

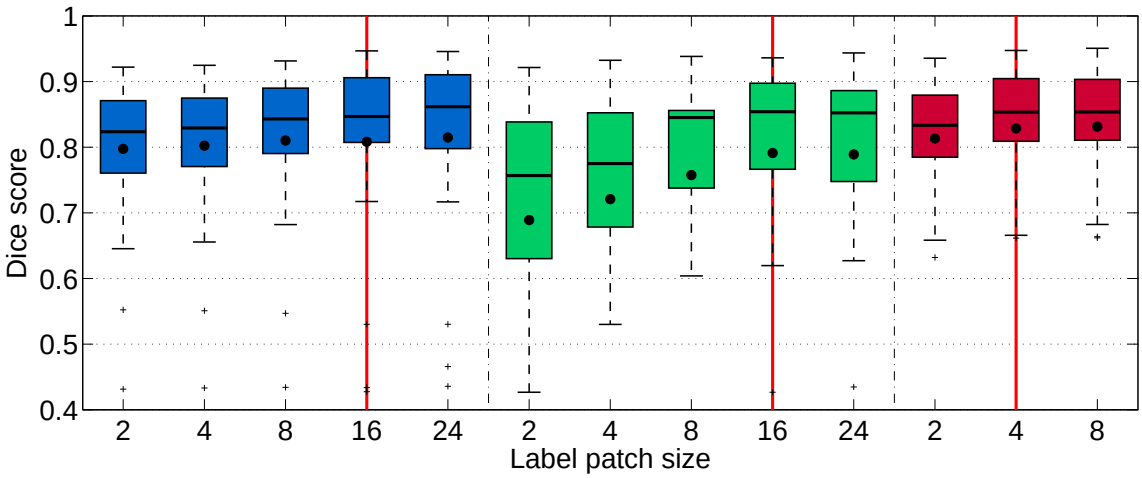


Fig. 6.7: Dice score as a function of the label patch size d' for whole tumor (blue) with $d = 24$, tumor core (green) with $d = 24$, and active tumor (red) with $d = 8$, with label patch dictionary size $N = 16$.

6.6.3 Application to the test set

After the optimization of the parameters using the validation set, the algorithm was tested on a new set of 66 subjects randomly chosen from BRATS 2014. The performance for both validation and test set of all three segmented structures is summarized in Tab. 6.22. For the test set, the achieved average Dice scores are 83% (whole tumor), 75% (tumor core), and 77% (active tumor).

Tab. 6.22: Segmentation results on validation and test data sets, reporting average and median Dice scores. Shown are the results for all three segmented structures, i.e., whole tumor, tumor core and active tumor. Scores for active tumor are calculated for high-grade cases only. "std" and "mad" denote standard deviation and median absolute deviance. HG and LG stand for high- and low-grade gliomas, respectively.

Dice Score (in %)	Whole	HG / LG	Core	HG / LG	Active
Validation set					
mean ± std	81±15	80±17 / 85±06	79±13	85±08 / 65±15	81±11
median ± mad	86±06	86±07 / 85±05	85±06	85±03 / 73±10	83±08
Test set					
mean ± std	83±13	86±09 / 76±21	75±20	79±14 / 61±29	77±18
median ± mad	88±04	88±03 / 87±05	83±08	82±07 / 72±14	83±09

Tables 6.23 and 6.24 summarize precision and sensitivity of the proposed algorithm for all three segmented structures. As it can be derived from the results in these two tables, the method has higher sensitivity than precision for the whole tumor region, which means that there are less false negatives than false positives, in other words the resulting area is usually larger than the true region. In average, 89% of the whole tumor area is labeled as pathological, while 81% of the resulting area is truly pathological. For tumor core and active tumor, the situation is opposite. High values of precision show that high percentage of voxels labeled as tumorous were selected as tumorous by expert too, i.e. 87% for tumor core and 85% for active tumor in average. However, the values of sensitivity are lower, especially for tumor core in case of low-grade gliomas. This means that there are less false positives than false negatives. In other words, the resulting area is smaller than the region selected manually by an expert.

Tab. 6.23: Segmentation results on validation and test data sets, reporting average and median precision. Shown are the results for all three segmented structures, i.e., whole tumor, tumor core and active tumor. Scores for active tumor are calculated for high-grade cases only. "std" and "mad" denote standard deviation and median absolute deviance. HG and LG stand for high- and low-grade gliomas, respectively.

Precision (in %)	Whole	HG / LG	Core	HG / LG	Active
Validation set					
mean ± std	78±16	77±18 / 81±07	92±06	91±07 / 93±03	90±10
median ± mad	84±08	83±08 / 84±08	93±03	93±03 / 92±01	92±04
Test set					
mean ± std	81±16	83±11 / 71±25	87±13	88±13 / 86±12	85±18
median ± mad	86±04	87±03 / 82±07	92±04	92±04 / 89±09	91±05

Examples of segmentations generated by the proposed method and corresponding manual segmentations for three segmented structures on representative test cases are shown in Fig. 6.8.

Compute time vs accuracy

The possibility of subsampling the volume in order to reduce the computational demands was tested here. The trade-off between accuracy and computing time per volume is analyzed in Tab. 6.25 by running several experiments with different resolutions of the CNN output before final prediction of the local structure (first column),

Tab. 6.24: Segmentation results on validation and test data sets, reporting average and median sensitivity. Shown are the results for all three segmented structures, i.e., whole tumor, tumor core and active tumor. Scores for active tumor are calculated for high-grade cases only. "std" and "mad" denote standard deviation and median absolute deviance. HG and LG stand for high- and low-grade gliomas, respectively..

Sensitivity (in %)	Whole	HG / LG	Core	HG / LG	Active
Validation set					
mean ± std	87±10	84±11 / 93±04	58±26	66±22 / 35±25	71±17
median ± mad	89±06	88±05 / 95±02	66±17	71±12 / 25±11	73±12
Test set					
mean ± std	89±09	89±09 / 89±08	66±24	71±21 / 47±27	72±23
median ± mad	92±03	92±02 / 90±06	71±16	78±11 / 53±17	80±12

i.e., subsampling in x and y, as well as different distances between segmented slices (second column), i.e., subsampling in z direction. All experiments were run on 4-core CPU Intel Xeon E3 3.30GHz. As one can see in the table, the state-of-the-art results can be achieved in an order of magnitude shorter time than in case of most methods participated in BRATS challenge. Thanks to the fast implementation of the CNN classification algorithm, all three structures can be segmented in the whole volume in 13 seconds without using GPU implementation. Processing by the CNN is approximately 80% of the overall computing time, while the assignment final labels using local structure prediction requires only 17%. The rest of the time are other operations including interpolation.

Discussion

In this section, it has been shown that exploiting local structure through the use of the label patch dictionaries improves segmentation performance over the standard approach predicting voxel wise labels. It has also been shown that local structure prediction can be combined with, and improves upon, standard prediction methods, such as CNNs. When optimized for a given segmentation problem it also performs a spatial regularization at the local level. On the reference benchmark set, the proposed approach achieved state-of-the-art performance even without post-processing through Markov random fields which were part of most best performing approaches in the tumor segmentation challenge. Moreover, all three structures can be extracted from the whole volume within only 13 seconds using CPU obtaining state-of-the-

Tab. 6.25: Trade-off between spatial subsampling, computing time, and segmentation accuracy. First two columns express different CNN output resolution, i.e., after subsampling in x and y, and steps between segmented slices, i.e., after subsampling in z direction.

CNN output resolution	Slice step	Computing time per volume	Dice Score (in%)		
			Whole	Core	Active
1/4	4	13s	83	75	73
1/4	2	22s	84	75	74
1/4	1	74s	84	75	75
1/2	4	24s	83	75	74
1/2	2	41s	83	75	76
1/2	1	142s	84	75	76
1/1	4	47s	83	75	75
1/1	2	80s	83	75	77
1/1	1	280s	83	75	77

art results providing means, for example, to do online updates when aiming at an interactive segmentation. The resulting Dice scores are comparable to intra-rater similarity that had been reported for the three annotation tasks in the BRATS data set [80] with Dice scores 85% (whole tumor), 75% (tumor core) and 74% (active tumor) and to the best results of automated segmentation algorithms with Dice scores of the top three in between 79%–82% (here: 83%) for the whole tumor segmentation task, 65%–70% (here: 75%) for the segmentation of the tumor core area, and 58%–61% (here: 77%) for the segmentation of the active tumor region.

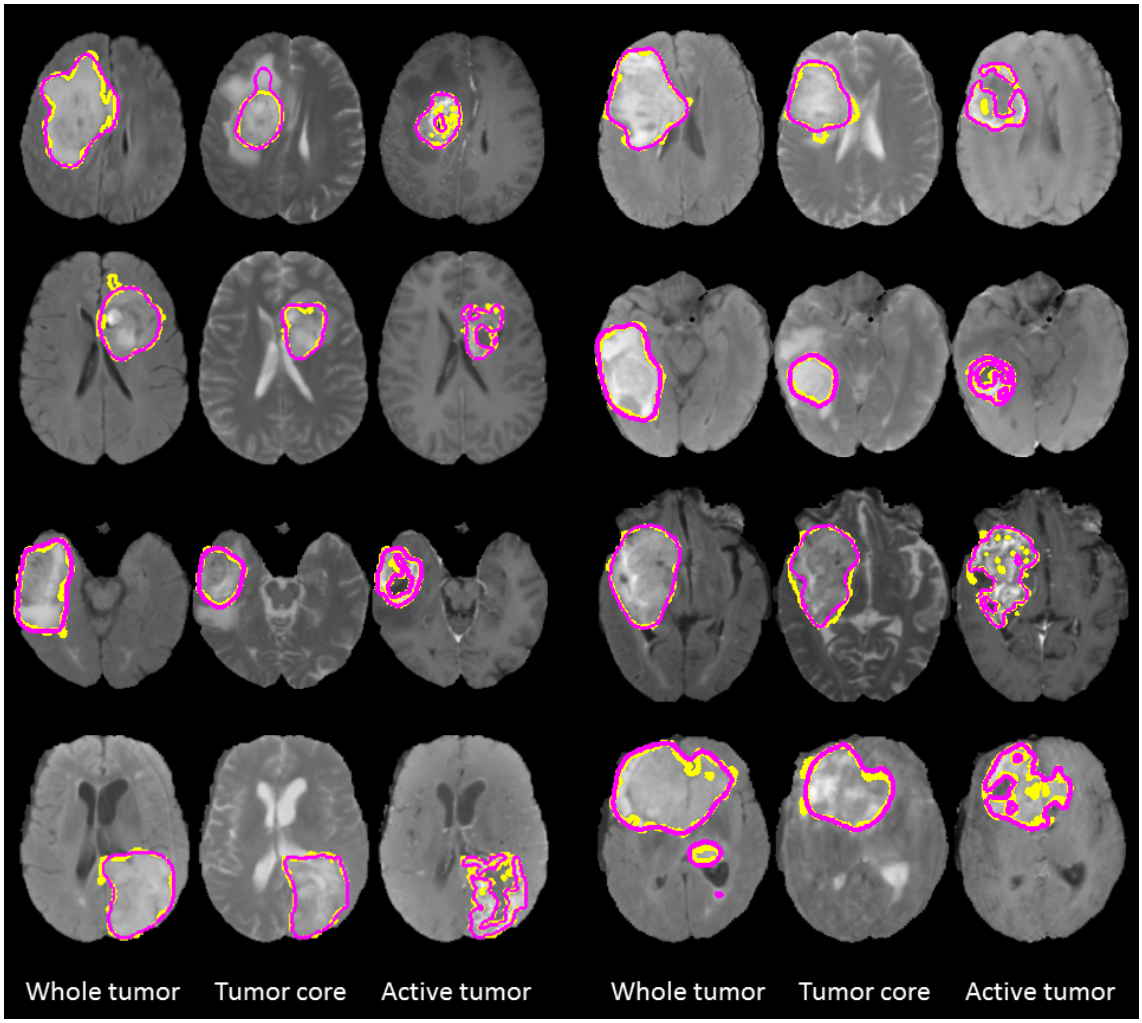


Fig. 6.8: Example of consensus expert annotation (yellow) and automatic segmentation (magenta) applied to the test image data set. Each row shows two cases. From left to right: segmentation of whole tumor (shown in FLAIR), tumor core (shown in T2) and active tumor (shown in T1c).

7 CONCLUSION

This thesis showed several approaches for brain tumor detection and segmentation in 2D or 3D single- and multisequence MRI. The attention was paid to high- and low-grade gliomas. Three algorithms dealing with this topic were introduced. All three algorithms were tested and evaluated on a large public BRATS challenge database of 254 3D multisequence subjects. It contained co-registered and skull-stripped FLAIR, T2, T1 and T1-contrast enhanced volumes with isotropic resolution 1mm. Therefore, the work did not deal with such pre-processing steps. All data sets included manual annotations provided by experts and they were used for the evaluation of the proposed methods.

The first two approaches explored the suitability of using prior brain anatomy knowledge to detect and extract brain tumors. The first method used this information for supervised detection of a particular brain tumor structure presence in 2D single- and multisequence images in both stand-alone slices and slices in 3D volume of planes where the left-right symmetry exists, i.e. axial and coronal. When applied slice-wise to 3D volumes, more information was used and higher accuracy was achieved. For stand-alone multisequence slices, where only the information extracted from asymmetry maps was used, the detection accuracy of each structure of size $\geq$1cm^2 was around 90%. When applied to the same slices with the use of asymmetry information together with image intensity information, which is only applicable in 3D volume, the accuracy increased to 93%. A slice was processed in only 0.11 seconds in average, which makes it suitable for either brain tumor locating in 3D volumes using the slice-wise prediction approach, or in a pipeline of any brain tumor segmentation techniques.

The second method applied similar methodology to locate a brain tumor in a singlesequence 2D and 3D MRI followed by its extraction from a multisequence MRI in an unsupervised manner. In 2D MR, the computing time of the whole extraction process was 0.13 seconds with reached average and median Dice score of 63$\pm$30 and 77$\pm$27, respectively. In 3D, the proposed method was able to locate the tumor in less than 8 seconds and extract it in about 13 seconds with reached average and median Dice score of 66$\pm$21 and 74$\pm$17, respectively. Compared to other state-of-the-art algorithms, both proposed methods are not influenced by the accuracy of the intensity normalization algorithm since they are independent on the intensity range.

The third method, focused on the segmentation of whole tumor, tumor core, and active tumor, showed that exploiting local structure through the use of the label patch dictionaries improved segmentation performance over the standard approach predicting voxel-wise labels. It was shown that local structure prediction

can be combined with, and improves upon, standard prediction methods, such as CNN. When the label patch size is optimized for a given segmentation task, it is capable of accumulating local evidence for a given label, and also performs a spatial regularization at the local level. The proposed approach achieved state-of-the-art performance even without sophisticated post-processing step which are part of most best performing approaches in brain tumor segmentation. Moreover, all three structures can be extracted from the whole volume within only 13 seconds using CPU obtaining state-of-the-art results providing means, for example, to do online updates when aiming at an interactive segmentation. Most medical image data consist of 3D volumes. Therefore, one of the possible future directions in the development of the local structure prediction algorithm can be an exploration of natural 3D implementation instead of slice-wise approach.

BIBLIOGRAPHY

[1] Q. Ain, I. Mehmood, S. M. Naqi, et al. "Bayesian Classification Using DCT Features for Brain Tumor Detection". In: *Proc KES*. 2010.

[2] A. Akselrod-Ballin, M. Galun, M. J. Gomori, et al. "An Integrated Segmentation and Classification Approach Applied to Multiple Sclerosis Analysis". In: *Proc CVPR*. 2006.

[3] R. D. Ambrosini, P. Wang, and W. G. O'Dell. "Computer-aided detection of metastatic brain tumors using automated three-dimensional template matching". In: *J. Magn. Reson. Imaging* 31 (2010), pp. 85–93.

[4] T. S. Armstrong, M. Z. Cohen, J. Weinbrg, et al. "Imaging techniques in neuro oncology". In: *Seminars in Oncology Nursing* 20 (2004), pp. 231–239.

[5] R. Ayachi and N. B. Amor. "Brain Tumor Segmentation Using Support Vector Machines". In: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. 2009, pp. 736–747.

[6] S. Bauer, L. P. Nolte, and M. Reyes. "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization". In: *Proc MICCAI*. 2011, pp. 354–361.

[7] A. Bhattacharyya. "On a measure of divergence between two statistical populations defined by their probability distribution". In: *Bulletin of the Calcutta Mathematical Society* 35 (1943), pp. 99–110.

[8] C. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[9] A. Bleau and L. Leon. "Watershed-based segmentation and region merging". In: *Comput Vis Image Under* 77(3) (2000), pp. 317–320.

[10] J. Bouvrie. *Notes on convolutional neural networks*. Tech. rep. MIT, 2006.

[11] A. Capelle, O. Alata, C. Fernandez, et al. "Unsupervised segmentation for automatic detection of brain tumors in MRI". In: *Proc ICIP 2000*. 2000.

[12] V. Caselles, F. Catte, T. Coll, et al. "A geometric model for active contours". In: *Numerische Mathematik*. Vol. 66. 1993, pp. 1–31.

[13] J. Cates, R. Whitaker, and J. G. "Case study: an evaluation of user-assisted hierarchical watershed segmentation". In: *Medical Image Analysis* (2005), pp. 566–578.

[14] J. E. Cates, A. E. Lefohn, and R. T. Whitaker. "GIST: An interactive GPUbased level-set segmentation tool for 3D medical images". In: *Med Image Anal* 8(3) (2004), pp. 217–231.

[15] S. Chaplot, L. M. Patnaik, and N. R. Jagannathan. "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network". In: *Biomedical Signal Processing and Control* 1 (2006), pp. 86–92.

[16] L.-C. Chen, G. Papandreou, and A. Yuille. "Learning a Dictionary of Shape Epitomes with Applications to Image Labeling". In: *Proc ICCV 2013*. 2013, pp. 337–344. DOI: 10.1109/ICCV.2013.49.

[17] M. Clark, L. Hall, D. Goldgof, et al. "Automatic tumor segmentation using knowledge-based techniques". In: *TMI* 17(2) (1998), pp. 187–201.

[18] L. Clarke. "MR image segmentation using MLM and artificial neural nets". In: *Med Phys* 18(30 (1991), pp. 672–682.

[19] D. Cobzas, N. Birkbeck, M. Schmidt, et al. "3D variational brain tumor segmentation using a high dimensional feature set". In: *Proc ICCV 2007*. 2007.

[20] J. J. Corso, E. Sharon, S. Dube, et al. "Efficient Multilevel Brain Tumor Segmentation with Integrated Bayesian Model Classification". In: *TMI* 27(5) (2011), pp. 629 –640.

[21] J. J. Corso, E. Sharon, and A. Yuille. "Multilevel Segmentation and Integrated Bayesian Model Classification with an Application to Brain Tumor Segmentation". In: *Medical Image Computing and Computer Assisted Intervention*. 2006, pp. 790–798.

[22] C. Cortes and V. Vapnik. "Support-Vector Networks". In: *Mach. Learn.* 20.3 (Sept. 1995), pp. 273–297. ISSN: 0885-6125. DOI: 10.1023/A:1022627411411. URL: http://dx.doi.org/10.1023/A:1022627411411.

[23] A. Criminisi and J. Shotton. *Decision Forests for Computer Vision and Medical Image Analysis*. Springer, 2013.

[24] E. Dam, M. Loog, and M. Letteboer. "Integrating automatic and interactive brain tumor segmentation." In: *17th International Conference on Pattern Recognition (ICPR 2004)*. Vol. 3. 2004, pp. 790–793.

[25] A. Davy, M. Havaei, D. Warde-Farley, et al. "Brain Tumor Segmentation with Deep Neural Networks". In: *Proc MICCAI-BRATS*. 2014.

[26] B. M. Dawant, S. L. Hartmann, S. Pan, et al. “Brain atlas deformation in the presence of small and large space-occupying tumors.” In: *Comput. Aided Surg.* 7(1) (2002), pp. 1–10.

[27] A. Dempster, N. Laird, and D. Rubin. “Maximum Likelihood from Incomplete Data via the EM Algorithm”. In: *Journal of the Royal Statistical Society* 39(1) (1977), pp. 1–38.

[28] S. Dickson and B. Thomas. “Using neural networks to automatically detect brain tumours in MR images”. In: *Int J Neural Syst* 4(1) (1997), pp. 91–99.

[29] P. Dollar and C. L. Zittnick. “Structured Forests for Fast Edge Detection”. In: *Proc ICCV 2013*. 2013, pp. 1841–1848.

[30] W. Dou, S. Ruan, Y. Chen, et al. “A framework of fuzzy information fusion for the segmentation of brain tumor tissues on {MR} images”. In: *Image and Vision Computing* 25.2 (2007). <ce:title>Soft Computing in Image Analysis</ce:title>, pp. 164 –171. ISSN: 0262-8856. DOI: http://dx.doi.org/10.1016/j.imavis.2006.01.025. URL: http://www.sciencedirect.com/science/article/pii/S0262885606000734.

[31] M. Droske, B. Meyer, M. Rumpf, et al. “An adaptive level set method for medical image segmentation”. In: *Proceedings IPMI*. 2001, pp. 416–422.

[32] E. Eisenhauer, P. Therasse, J. Bogaerts, et al. “New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)”. In: *European Journal of Cancer* 45 (2) (2009), pp. 228–247.

[33] R. Farjam, H. A. Parmar, D. C. Noll, et al. “An approach for computer-aided detection of brain metastases in post-Gd T1-W MRI”. In: *Magn. Reson. Imaging* 30 (2012), pp. 824–836.

[34] C. Fennema-Notestine, I. B. Ozyurt, C. P. Clark, et al. “Quantitative Evaluation of Automated Skull-Stripping Methods Applied to Contemporary and Legacy Images: Effects of Diagnosis, Bias Correction, and Slice Location”. In: *Human Brain Mapping* 27(2) (2006), pp. 99–113.

[35] J. Festa, S. Pereira, J. A. Mariz, et al. “Automatic Brain Tumor Segmentation of Multi-sequence MR images using Random Decision Forests MICCAI Grand Challenge: BRATS 2013”. In: *Proc MICCAI-BRATS*. 2013.

[36] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, et al. “Automatic segmentation of non-enhancing brain tumor in magnetic resonance images.” In: *Artificial Intelligence in Medicine* 21 (2001), pp. 43–63.

[37] V. S. Fonov, A. C. Evans, K. Botteron, et al. “Unbiased average age-appropriate atlases for pediatric studies”. In: *NeuroImage* 54(1) (2011).

[38] B. J. Frey and D. Dueck. “Clustering by Passing Messages Between Data Points”. In: *Science* 315 (2007), pp. 972–976. URL: www.psi.toronto.edu/affinitypropagation.

[39] C. Garcia and J. Moreno. “Kernel based method for segmentation and modeling of magnetic resonance images”. In: *LNCS, 3315*. 2004, pp. 636–645.

[40] E. Geremia, O. Clatz, B. H. Menze, et al. “Spatial Decision Forests for MS Lesion Segmentation in Multi-Channel Magnetic Resonance Images”. In: *Neuroimage* 57(2) (2011), pp. 378–390.

[41] E. Geremia, D. Zikic, O. Clatz, et al. “Decision Forests for Computer Vision and Medical Image Analysis Advances in Computer Vision and Pattern Recognition”. In: Springer, 2013. Chap. Classification Forests for Semantic Segmentation of Brain Lesions in Multi-channel MRI, pp. 245–260.

[42] E. Geremia, B. H. Menze, and N. Ayache. “SPATIALLY ADAPTIVE RANDOM FORESTS”. In: *Proc ISBI*. 2013.

[43] D. T. Gering. “Recognizing Deviations from Normalcy for Brain Tumor Segmentation”. PhD thesis. Massachusetts Institute of Technology, 2003.

[44] P Gibbs, D. L. Buckley, S. J. Blackband, et al. “Tumour volume determination from MR images by morphological segmentation”. In: *Physics in Medicine and Biology* 41 (1996), pp. 2437–2446.

[45] M. Goetz, C. Weber, J. Bloecher, et al. “Extremely randomized trees based brain tumor segmentation”. In: *Proc MICCAI-BRATS 2014*. 2014.

[46] M. Havaei, A. Davy, D. Warde-Farley, et al. “Brain Tumor Segmentation with Deep Neural Networks”. In: (2015), pp. 1–29.

[47] M. Havaei, P.-M. Jodoin, and H. Larochelle. “Efficient interactive brain tumor segmentation as within-brain kNN classification”. In: *Proc ICPR*. 2014.

[48] S. Ho, E. Bullitt, and G. Gerig. “Level-Set Evolution with Region Competition: Automatic 3-D Segmentation of Brain Tumors”. In: *Proc ICPR*. ICPR ’02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 10532–. ISBN: 0-7695-1695-X. URL: http://dl.acm.org/citation.cfm?id=839290.842671.

[49] C.-W. Hsu and C.-J. Lin. “A comparison of methods for multiclass support vector machines”. In: *Neural Networks, IEEE Transactions on* 13(2) (2002), pp. 415–425.

[50] J. E. Iglesias, E. Konukoglu, D. Zikic, et al. “Is Synthesizing MRI Contrast Useful for Inter-modality Analysis?” In: *Proc MICCAI 2013*. 2013, pp. 631–638.

[51] A. Islam, S. M. S. Reza, and K. M. Iftekharuddin. “Multifractal Texture Estimation for Detection and Segmentation of Brain Tumors”. In: *IEEE Trans Biomed Eng.* 60 (11) (2013), pp. 3204–3215.

[52] M. Kaleem, M. Sanaullah, M. A. Hussain, et al. “Segmentation of Brain Tumor Tissue Using Marker Controlled Watershed Transform Method”. In: *Emerging Trends and Applications in Information Communication Technologies*. Vol. 281. 2012, pp. 222–227.

[53] M. R. Kaus, S. K. Warfield, A. Nabavi, et al. “Segmentation of meningiomas and low grade gliomas in MRI”. In: *Proc MICCAI*. 1999, pp. 1–10.

[54] M. R. Kaus, S. K. Warfield, A. Nabavi, et al. “Automated Segmentation of MR Images of Brain Tumors”. In: *Radiology* 218(2) (2001), pp. 586–591.

[55] N. E. A. Khalid, S. Ibrahim, and P. N. M. M. Haniff. “MRI Brain Abnormalities Segmentation usingK-Nearest Neighbors”. In: *International Journal on Computer Science and Engineering* 3(2) (2011), pp. 980–990.

[56] O. Khotanlou H. Colliot and I. Bloch. “Automatic brain tumor segmentation using symmetry analysis and deformable models”. In: *Conf. on Advances in Pattern Recognition ICAPR, Kolkata, India,* (2007).

[57] J. Kong, J. Wang, Y. Lu, et al. “. A novel approach for segmentation of MRI brain image”. In: *IEEE MELECON* (2006), pp. 525–528.

[58] P. Kontschieder, S. Rota Bulo, H. Bischof, et al. “Structured class-labels in random forests for semantic image labelling”. In: *Proc ICCV 2011*. 2011, pp. 2190–2197. DOI: 10.1109/ICCV.2011.6126496.

[59] P. Kontschieder, S. R. Bulo, M. Donoser, et al. “Semantic Image Labelling as a Label Puzzle Game”. In: *Proceedings of the British Machine Vision Conference (BMVC)*. 2011, pp. 111.1–111.12. DOI: http://dx.doi.org/10.5244/C.25.111.

[60] A. Krizhevski, I. Sutskever, and G. E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *NIPS 2012*. 2012.

[61] D Kumar and A Halder. “Automatic brain tumor detection and isolation of tumor cells from MRI images”. In: *Int J Comput Appl* 39(16) (2012), pp. 26–30.

[62] S. K. Kyriacou, C. Davatzikos, S. J. Zinreich, et al. “Nonlinear Elastic Registration of Brain Images with Tumor Pathology Using a Biomechanical Model”. In: *TMI* 18(7) (1999), pp. 580–592.

[63] S. Lakare. *3D segmentation techniques for medical volumes. , state university of NY; 2000.* Stony Brooks: Center of Visual Computer, State university of NY, 200.

[64] A. Law, F. Lam, and F. Chan. "A fast deformable region model for brain tumor boundary extraction". In: *Int Conf EMBS*. Vol. 2. 2002, pp. 1055–1056.

[65] Y. LeCun, B. Boser, J. Denker, et al. "Back-propagation applied to handwritten zip code recognition". In: *Neural Computation* 1(4) (1989).

[66] Y. LeCun, B. Boser, J. S. Denker, et al. "Advances in neural information processing systems 2". In: Morgan Kaufmann Publishers Inc., 1990. Chap. Handwritten digit recognition with a back-propagation network, pp. 396–404.

[67] Y. LeCun and Y. Bengio. "The Handbook of Brain Theory and Neural Networks". In: ed. by M. A. Arbib. Cambridge, MA, USA: MIT Press, 1998. Chap. Convolutional Networks for Images, Speech, and Time Series, pp. 255–258. ISBN: 0-262-51102-9. URL: `http://dl.acm.org/citation.cfm?id=303568.303704`.

[68] C. Lee, S. Wang, A. Murtha, et al. "Segmenting brain tumors using pseudo-conditional random fields". In: *Proc MICCAI 2008*. 2008.

[69] A. Lefohn, J. Cates, and R. Whitaker. *Interactive GPU-based level sets for 3D brain tumor segmentation*. Tech. rep. University of Utah, 2003.

[70] M. Letteboer, O. Olsen, E. Dam, et al. "Segmentation of tumors in magnetic resonance brain images using an interactive multiscale watershed algorithm". In: *Academic Radiology* 11(10) (2004), pp. 1125–1138.

[71] S. Liao, Y. Gao, A. Oto, et al. "Representation Learning: A Unified Deep Learning Framework for Automatic Prostate MR Segmentation". In: *Proc MICCAI 2013*. 2013, pp. 254–261.

[72] Y. Liu, R. T. Collins, and W. E. Rothfus. "Robust midsagittal plane extraction from normal and pathological 3-D neuroradiology image". In: *IEEE Transactions on Medical Imaging* 20(3) (2003), pp. 175–192.

[73] S. Luo. "Automated medical image segmentation using a new deformable surface model". In: *Int J Comput Sci Netw Secur* 6(5A) (2006), pp. 109–115.

[74] S. Luo, R. Li, and S. Ourselin. "A new deformable model using dynamic gradient vector flow and adaptive balloon forces". In: *APRS Workshop on Dig Image Comp*. 2003, pp. 9–14.

[75] M. Mancas and B. Gosselin. “Toward an automatic tumor segmentation using iterative watersheds”. In: *Medical Imaging 2004: Image Processing (SPIE2004)*. Vol. 5370. 2004, pp. 1598–1608.

[76] D. Mattes, D. R. Haynor, H. Vesselle, et al. “Nonrigid multimodality image registration”. In: *Medical Imaging 2001: Image Processing* (2001), pp. 1609–1620.

[77] T. McInerney and D. Terzopoulos. “Deformable models in medical image analysis: a survey”. In: *Med. Image. Anal.* 1 (1996), pp. 91–108.

[78] R. Meier, S. Bauer, J. Slotboom, et al. “A Hybrid Model for Multimodal Brain Tumor Segmentation”. In: *Proc MICCAI-BRATS 2013*. 2013.

[79] R. Meier, S. Bauer, J. Slotboom, et al. “Appearance- and Context-sensitive Features for Brain Tumor Segmentation”. In: *Proceedings MICCAI-BRATS 2014*. 2014.

[80] B. Menze, A. Jakab, S. Bauer, et al. “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”. In: *IEEE Transactions on Medical Imaging* (2014), p. 33. DOI: 10.1109/TMI.2014.2377694. URL: https://hal.inria.fr/hal-00935640.

[81] B. Menze, K. van Leemput, D. Lashkari, et al. “A Generative Model for Brain Tumor Segmentation in Multi-Modal Images”. English. In: *Proc MICCAI 2010*. 2010, pp. 151–159. ISBN: 978-3-642-15744-8. DOI: 10.1007/978-3-642-15745-5_19. URL: http://dx.doi.org/10.1007/978-3-642-15745-5_19.

[82] B. H. Menze, K. V. Leemput, A. Honkela, et al. “A Generative Approach for Image-Based Modeling of Tumor Growth”. In: *Proc MICCAI 2011*. 2011, pp. 735–747.

[83] J. Mikulka and E. Gescheidtova. “An Improved Segmentation of Brain Tumor, Edema and Necrosis”. In: *Progress in Electromagnetics Research Symposium*. 2013, pp. 25–28.

[84] N. Moon, E. Bullitt, K. van Leemput, et al. “Model-Based Brain and Tumor Segmentation”. In: *Proc ICPR 2002*. 2002.

[85] S. Murugavalli and V. Rajamani. “An improved implementation of brain tumor detection using segmentation based on neuro fuzzy technique”. In: *J Comput Sci* 3(110 (2007), pp. 841–846.

[86] N. Otsu. “A Threshold Selection Method from Gray-Level Histograms”. In: *IEEE Transactions on Systems, Man and Cybernetics* 9.1 (1979), pp. 62–66. DOI: 10.1109/TSMC.1979.4310076. URL: http://dx.doi.org/10.1109/TSMC.1979.4310076.

[87] M. Ozkan, B. Dawant, and R. Maciunas. “Neural-network-based segmentation of multi-modal medical images: a comparative and prospective study”. In: *TMI* 12(3) (1993), pp. 534–544.

[88] W. Phillips, R. Velthuizen, S. Phupanich, et al. “Application of fuzzy C-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme.” In: *Magn Reson Imag* 13(20 (1995), pp. 277–290.

[89] P. H. O. Pinheiro and R. Collobert. “Recurrent Convolutional Neural Networks for Scene Labeling”. In: *International Conference on Machine Learning (ICML)*. 2014.

[90] A. Pinto, S. Pereira, H. Dinis, et al. “Random decision forests for automatic brain tumor segmentation on multi-modal MRI images”. In: *Proc ENBENG 2015*. 2015.

[91] K. M. Pohl, S. Bouix, M. Nakamura, et al. “A Hierarchical Algorithm for MR Brain Image Parcellation”. In: *TMI* 26(9) (2007), pp. 1201–1212.

[92] K. Popuri, D. Cobzas, M. Jagersand, et al. “3D variational brain tumor segmentation on a clustered feature set”. In: *SPIE Med Imaging*. 2009.

[93] A. Prasoon, K. Petersen, C. Igel, et al. “Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network”. In: *MICCAI*. 2013, pp. 246–253.

[94] M. Prastawa, E. Bullitt, S. Ho, et al. “A brain tumor segmentation framework based on outlier detection”. In: *Med Image Anal* 8 (2004), pp. 275–283.

[95] M. Prastawa, E. Bullitt, N. Moon, et al. “Automatic Brain Tumor Segmentation by Subject Specific Modification of Atlas Priors”. In: *Academic Radiology* 10 (12) (2003), pp. 1341–1348.

[96] F. Prior, S. J. Fouke, T. Benzinger, et al. “Predicting a multi-parametric probability map of active tumor extent using random forests”. In: *Proc EMBC 2013*. 2013.

[97] W. Reddick, J. Glass, E. Cook, et al. “Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks”. In: *TMI* 16(6) (1997), pp. 911–918.

[98] J Rexilius, H Hahn, J Klein, et al. “Multispectral brain tumor segmentation based on histogram model adaptation”. In: *SPIE Conf Med Image Comp*. 2007.

[99] S. Reza and K. M. Iftekharuddin. “Multi-class Abnormal Brain Tissue Segmentation Using Texture Features”. In: *Proc MICCAI-BRATS 2013*. 2013.

[100] S. Reza and K. M. Iftekharuddin. “Multi-fractal texture features for brain tumor and edema segmentation”. In: *Proc SPIE Medical Imaging 2014*. 2014.

[101] H. R. Roth, L. Lu, A. Seff, et al. “A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations”. In: *MICCAI 2014, Part I*. 2014, pp. 520–527.

[102] S. Ruan, S. Lebonvallet, A. Merabet, et al. “Tumor segmentation from a multispectral MRI images by using support vector machine classification”. In: *Proc ISBI*. 2007, pp. 1236–1239.

[103] G. C. S. Ruppert, L. Teverovskiy, C. Yu, et al. “A New Symmetry-based Method for Mid-sagittal Plane Extraction in Neuroimages.” In: *International Symposium on Biomedical Imaging: From Macro to Nano*. 2011.

[104] B. N. Saha, N. Ray, R. Greiner, et al. “Quick detection of brain tumors and edemas: A bounding box method using symmetry”. In: *Computerized Medical Imaging and Graphics* 36.2 (2012), pp. 95 –107. ISSN: 0895-6111. DOI: http://dx.doi.org/10.1016/j.compmedimag.2011.06.001. URL: http://www.sciencedirect.com/science/article/pii/S0895611111000796.

[105] Y Salman. “Modified technique for volumetric brain tumor measurements”. In: *J Biomed Sci Eng* 2 (2009), pp. 16–19.

[106] Y. Salman and A. Badawi. “Validation techniques for quantitative brain tumor measurements”. In: *Int Conf IEEE EMBS*. 2005, pp. 7048–7051.

[107] Y. Salman, A. Badawi, M. Assal, et al. “New automatic technique for tracking brain tumor response”. In: 2005, pp. 1–4.

[108] M Sato, S Lakare, M Wan, et al. “A gradient magnitude based region growing algorithm for accurate segmentation”. In: *Int Conf Image Process*. 2000, pp. 448–451.

[109] K. Shanthi and M. Kumar. “Skull stripping and automatic segmentation of brain MRI using seed growth and threshold techniques”. In: *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*. 2007, pp. 422–426. DOI: 10.1109/ICIAS.2007.4658421.

[110] E. Sharon, A. Brandt, and R. Basri. “Segmentation and boundary detection using multiscale intensity measurements”. In: *Proc CVPR*. 2001, pp. 469–476.

[111] S. Shen, W. A. Sandham, and M. H. Granat. “Preprocessing and segmentation of brain magnetic resonance images.” In: *IEEE Conf. on Information Technology Applications in Biomedicine*. 2003, pp. 149–152.

[112] F. Y. Shih and S. Cheng. “Automatic seeded region growing for color image segmentation”. In: *Image and Vision Computing* 23.10 (2005), pp. 877 –886. ISSN: 0262-8856. DOI: http://dx.doi.org/10.1016/j.imavis.2005.05.015. URL: http://www.sciencedirect.com/science/article/pii/S0262885605000673.

[113] P. P. Singhai and S. A. Ladhake. “Brain Tumor Detection Using Marker Based Watershed Segmentation from Digital MR Images”. In: *International Journal of Innovative Technology and Exploring Engineering* 2(5) (2013), pp. 201–204.

[114] J. Solomon, J. A. Butman, and A. Sood. “Segmentation of brain tumors in 4D MR images using the hidden Markov model”. In: *Computer Methods and Programs in Biomedicine* 84 (2006), pp. 76–85.

[115] A. Stadlbauer, E. Moser, S. Gruber, et al. “Improved delineation of brain tumors: an automated method for segmentation based on pathologic changes of 1H-MRSI metabolites in gliomas”. In: *NeuroImage* 23.2 (2004), pp. 454 –461. ISSN: 1053-8119. DOI: http://dx.doi.org/10.1016/j.neuroimage.2004.06.022. URL: http://www.sciencedirect.com/science/article/pii/S1053811904003313.

[116] R. Stupp, W. Mason, M. van den Bent, et al. “Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma”. In: *New England Journal of Medicine* 352(10) (2005), pp. 987–996.

[117] L. Szilagyi, S. Szilagyi, and Z. Benyo. “A modified fuzzy C-means algorithm for MR brain image segmentation”. In: *Image Analysis and Recognition*. 2007.

[118] S. Taheri, S. Ong, and V. Chong. “Level-set segmentation of brain tumors using a threshold-based speed function”. In: *Image and Vision Computing* 28(1) (2010), pp. 26–37.

[119] K. Thapaliya, J.-Y. Pyun, C.-S. Park, et al. “Level set method with automatic selective local statistics for brain tumor segmentation in MR images”. In: *Computerized Medical Imaging and Graphics* 37 (2013), pp. 522–537.

[120] T. Tong, R. Wolz, P. Coupe, et al. “Segmentation of MR images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling”. In: *NeuroImage* 76 (2013), pp. 11–23.

[121] N. Tustison, B. Avants, P. Cook, et al. “N4ITK: Improved N3 Bias Correction with Robust B-Spline Approximation”. In: *Proc. of ISBI*. 2010.

[122] G. Urban, M. Bendszus, F. Hamprecht, et al. “Multi-modal brain tumor segmentation using deep convolutional neural networks”. In: *Proc MICCAI-BRATS*. 2014, pp. 31–35.

[123] C. Vijayakumar, G. Damayanti, R. Pant, et al. “Segmentation and grading of brain tumors on apparent diffusion coefficient images using self-organizing maps”. In: *Comput Med Imaging Graph* 31 (2007), pp. 473–484.

[124] F. Vincent, H. Chong, J. Zhou, et al. “Tongue carcinoma: tumor volume measurement.” In: *International Journal of Radiation Oncology, Biology, Physics* 59(1) (2004), pp. 59–66.

[125] S. Vinitski, T. Iwanaga, C. Gonzalez, et al. “Fast tissue segmentation based on a 4D feature map: Preliminary results”. In: *Proc ICIAP*. 1997, pp. 445–452.

[126] T. Wang, I. Cheng, and A. Basu. “Fully automatic brain tumor segmentation using a normalized Gaussian Bayesian Classifier and 3D Fluid Vector Flow”. In: *Proc ICIP*. 2010, pp. 2553–2556.

[127] T. Weglinski and A. Fabijanska. “Brain tumor segmentation from MRI data sets using region growing approach”. In: *Perspective Technologies and Methods in MEMS Design (MEMSTECH), 2011 Proceedings of VIIth International Conference on*. 2011, pp. 185–188.

[128] C. Xu, D. L. Pham, and J. L. Prince. “Handbook of Medical Imaging”. In: SPIE Press, 2000. Chap. Medical Image Segmentation Using Deformable Models: Medical Image Processing and Analysis, pp. 129–174.

[129] J. G. Zhang, K. K. Ma, M. H. Er, et al. “Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine”. In: *Proc IWAIT 2004*. 2004.

[130] N. Zhang. “Feature Selection based Segmentation of Multi-Source Images: Application to Brain Tumor Segmentation in Multi-Sequence MRI”. PhD thesis. INSA de Lyon, 2011.

[131] N. Zhang, S. Ruan, S. Lebonvallet, et al. “Multi-kernel SVM based classification for brain tumor segmentation of MRI multi-sequence”. In: *Proc ICIP*. 2009.

[132] Y. Zhang, M. Brady, and S. Smith. “Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm”. In: *TMI* 20(1) (2001), pp. 45–57.

[133] L. Zhao, W. Wu, and J. J.Corso. “Brain tumor segmentation based on GMM and active contour method with a model-aware edge map”. In: *Proc BRATS 2012*. 2012.

[134] L. Zhu, Y. Chen, Y. Lin, et al. “Recursive Segmentation and Recognition Templates for 2D Parsing”. In: *NIPS*. 2009, pp. 1985–1992.

[135] L. Zhu, Y. Chen, Y. Lin, et al. “Recursive segmentation and recognition templates for image parsing”. In: *IEEE PAMI* 34.2 (2012), pp. 359–371. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2011.160.

[136] D. Zikic, B. Glocker, E. Konukoglu, et al. “Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR”. In: *Proc MICCAI*. 2012.

[137] D. Zikic, Y. Ioannou, M. Brown, et al. “Segmentation of Brain Tumor Tissues with Convolutional Neural Networks”. In: *Proceedings MICCAI-BRATS 2014*. 2014, pp. 36–39.

BIBLIOGRAPHY OF THE AUTHOR

[138] R. Benes, P. Dvorak, M. Faundez-Zanuy, et al. "Multi- focus thermal image fusion". In: *Pattern Recognition Letters* 34(5) (2013), pp. 536–544.

[139] P. Dvorak. "Object description in images". In: *Student EEICT Proceedings of the 17th Conference Student EEICT 2010 Volume 2*. 2011, pp. 76–78.

[140] P. Dvorak and K. Bartusek. "Automated 3D Brain Tumor Edema Segmentation in FLAIR MRI". In: *MAGMA - ESMRMB 2013*. 2013.

[141] P. Dvorak and K. Bartusek. "Brain tumor locating in 3D MR volume using symmetry". In: *Proc. SPIE 9034, Medical Imaging 2014: Image Processing, 903432*. 2014. DOI: 10.1117/12.2042845.

[142] P. Dvorak and K. Bartusek. "Fully Automatic 3D Glioma Extraction in Multi- contrast MRI". In: *International Conference on Image Analysis and Recognition*. 2014, pp. 239–246.

[143] P. Dvorak and K. Bartusek. "Multi-parametric segmentation of MR images of the Brain". In: *9th International Conference on Measurement*. 2013, pp. 125–128.

[144] P. Dvorak, K. Bartusek, and E. Gescheidtova. "Automatic Extraction of Pathological Area in 2D MR Brain Scan". In: *PIERS 2014*. 2014, pp. 1885–1889.

[145] P. Dvorak, K. Bartusek, and E. Gescheidtova. "Automatic Segmentation of Multi- Contrast MRI Using Statistical Region Merging". In: *PIERS 2014*. 2014, pp. 1865–1869.

[146] P. Dvorak, K. Bartusek, and W. G. Kropatsch. "Automated Segmentation of Brain Tumor Edema in FLAIR MRI Using Symmetry and Thresholding". In: *Progress in Electromagnetics Research Symposium*. 2013, pp. 936–939.

[147] P. Dvorak, K. Bartusek, and W. G. Kropatsch. "Pathological Area Detection in MR Images of Brain". In: *Elektrorevue* 4(1) (2013), pp. 17–21.

[148] P. Dvorak, K. Bartusek, W. G. Kropatsch, et al. "Automated Multi- Contrast Brain Pathological Area Extraction from MR Images". In: *Journal of Applied Research and Technology* 13(1) (2015), pp. 58–69.

[149] P. Dvorak, K. Bartusek, and Z. Smekal. "Unsupervised Pathological Area Extraction Using 3D T2 and FLAIR MR Images". In: *Measurement Science Review* 14(6) (2014), pp. 357–364.

[150] P. Dvorak, V. Juras, and W. Vogl. "Interactive Hip Joint Cartilage Segmentation". In: *PIERS 2014*. 2014, p. 1563.

[151] P. Dvorak, V. Juras, W. Vogl, et al. “Interactive segmentation of Hip Joint Cartilage”. In: *PIERS 2014*. 2014, pp. 2369–2372.

[152] P. Dvorak, W. Kropatsch, and K. Bartusek. “Automatic detection of brain tumors in MR images”. In: *Telecommunications and Signal Processing (TSP), 2013 36th International Conference on*. 2013, pp. 577–580. DOI: 10.1109/TSP.2013.6614000.

[153] P. Dvorak and W. G. Kropatsch. “Detection of Brain Tumors Based on Automatic Symmetry Analysis”. In: *Proceedings of the 18th Computer Vision Winter Workshop*. 2013, pp. 24–31.

[154] P. Dvorak, W. G. Kropatsch, and K. Bartusek. “Automatic Brain Tumor Detection in T2-weighted Magnetic Resonance Images”. In: *Measurement Science Review* 13(5) (2013), pp. 223–230.

[155] P. Dvorak and B. H. Menze. “Structured Prediction with Convolutional Neural Networks for Multimodal Brain Tumor Segmentation”. In: *Proceedings MICCAI-MCV 2015*. 2015 (accepted).

[156] P. Dvorak, J. Mikulka, and K. Bartusek. “Tissue Segmentation of Brain MRI”. In: *37th International Conference on Telecommunications and Signal Processing 2014*. 2014, pp. 482–485.

[157] P. Dvorak and M. Zukal. “Description of Objects in Images Using MPEG- 7 Descriptors”. In: *Eletrorevue* 3(1) (2012), pp. 26–31.

[158] J. Mikulka and P. Dvorak. “Fast Calculation of T2 Relaxation Time in Magnetic Resonance Imaging”. In: *PIERS 2014*. 2014, pp. 2331–2335.

[159] J. Mikulka, P. Dvorak, and K. Bartusek. “Support Vector Machines in MR Images Segmentation”. In: *9th International Conference on Measurement*. 2013, pp. 157–160.

[160] M. Zukal, P. Dvorak, and P. Cika. “Implementation of MPEG-7 Descriptors”. In: *13th International Conference on Research in Telecommunication Technologies 2011*. 2011, pp. 127–130.

LIST ABBREVIATIONS

ANN	Artificial Neural Network
ADC	Apparent Diffusion Coefficient
AC	Asymmetry Coefficient
BC	Bhattacharyya Coefficient
BRATS	Brain Tumor Segmentation
CBTRUS	Central Brain Tumor Registry of the United State
CNN	Convolutional Neural Network
CNS	Central Nervous System
CPU	Central Processing Unit
CRF	Conditional Random Field
CSF	Cerebrospinal Fluid
DL	Deep Learning
DBN	Deep Belief Network
DBM	Deep Boltzmann Machine
DS	Dice Score
DW	Diffusion Weighted
EM	Expectation Maximization
FCM	Fuzzy C-Means
FLAIR	Fluid-Attenuated Inversion Recovery
GBM	Glioblastoma Multiform
GM	Gray Matter
GMM	Gaussian Mixture Model
HG	High Grade
HIM	Hierarchical Image Models

HOG	Histogram of Oriented Gradients
ITK	Insight Segmentation and Registration Toolkit
k-NN	K-Nearest Neighbors
LG	Low Grade
MAP	Maximum a posteriori
MR	Magnetic Resonance
MRF	Markov Random Field
MR	Magnetic Resonance Imaging
MLP	Multi-Layer Perceptron
NN	Neural Network
Pr	Probability
PD	Proton Density
PDF	Probability Density Function
PPV	Positive Predictive Value
RF	Random Forest
SAE	Stacked Auto-Encoder
SeCRF	Shape epitome Conditional Random Field
SOM	Self-Organizing Map
SVM	Support Vector Machine
SWI	Susceptibility Weighted Imaging
T1	T1-weighted MRI
T1C	Contrast enhanced T1-weighted MRI
T2	T2-weighted MRI
WM	White Matter

Curriculum Vitae

Pavel Dvořák

Personal information:

Address:	Přísnotice, Czech Republic
E-mail:	pavel.dvorak@hotmail.com
GSM:	+420 777 057 705

School and education:

09/2011-present	Brno University of Technology, Faculty of Electrical Engineering and Communication Ph.D.
09/2009-06/2011	Brno University of Technology, Faculty of Electrical Engineering and Communication MSc.
07/2004-06/2007	Brno University of Technology, Faculty of Electrical Engineering and Communication BSc.

Work experience:

01/2013-present	Junior Researcher, Brno University of Technology
02/2013-present	Junior Researcher, Czech Academy of Sciences

Research stays:

07/2015-08/2015	Escola Universitária Politécnica de Mataró, Spain
09/2014-03/2015	Chair for Computer Aided Medical Procedures & Augmented Reality, Technical University Munich, Germany
03/2014-08/2014	Private University of Health Sciences, Medical Informatics and Technology, Austria
12/2013-01/2014	Center of Medical Physics and Biomedical Engineering, Medical University of Vienna, Austria
09/2013-11/2013	MR Centre of Excellence, Medical University of Vienna, Austria
09/2012-12/2012	Pattern Recognition and Image Processing Group, Vienna University of Technology, Austria
06/2012-07/2012	Escola Universitária Politécnica de Mataró, Spain
11/2011-12/2011	Escola Universitária Politécnica de Mataró, Spain

Project participation:

04/2014-present	LD14091, Primary and secondary biometrical and non-biometrical features in the process of de-identification of audio-visual content in multimedia data
01/2014-present	FEKT-S-14-2335: Cognitive multimedia analysis of audio and image signals
02/2013–12/2014	GAP102/12/1104, Study of metabolism and localization of primary brain tumors using MR imaging techniques
02/2013–12/2013	GAP102/12/1104, Study of metabolism and localization of primary brain tumors using MR imaging techniques
01/2013-12/2013	ED2.1.00/03.0072, Center of Sensor, Information and Communication Systems
01/2012-12/2014	FEKT-S-11-17: Research of Sophisticated Methods for Digital Audio and Image Signal Processing

Awards:

2011	Dean’s award for the master’s thesis Object Description in Images
2011	Master Degree with Distinction

Publications:

- Papers published in journals with impact factor: 4
- Papers published in other journals: 2
- Papers published in international conference proceedings: 13
- Papers published in domestic conference proceedings: 1
- Abstracts published in international conference proceedings: 2
- Papers indexed in WoS: 5
- Papers indexed in Scopus: 11
- H-index according to WoS: 1
- H-index according to Scopus: 3