SAPIENZA
UNIVERSITÀ DI ROMA

My Thesis

Facoltà d'ingegneria dell'informazione, informatica e statistica
Computer Science - Informatica

Mattia Capparella
ID number 1746513

Advisor
Prof. de Marsico Maria

External advisor
Orlandi Manuel

Academic Year 2020/2021

Thesis not yet defended

My Thesis
Tesi di Laurea Magistrale. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: capparella.1746513@studenti.uniroma1.it

To the infinity and beyond

Abstract

Contents

Chapter 1

Introduction

Image segmentation (IS) is the process of partitioning an image into multiple image segments or objects. The goal of segmentation is to simplify the representation of an image into something that is more meaningful and easier to analyze. IS is a critical task in medical image analysis: it is often the first step to transform raw biomedical image into structured, valuable information ready to be used both for scientific discoveries and clinical applications including early diagnosis during preclinical phase, therapy planning, intraoperative assistance and tumor growth monitoring. Brain tumor segmentation is the process of isolating the tumor from healthy brain tissue; however, it is still a challenging task due to the irregular form and confusing boundaries of tumors.

Simply put, one major challenge is the lack of open datasets for designing and testing new algorithms, while private datasets may differ for so many aspects that comparing the result obtained by different solutions has no relevance and it's often inconclusive. On the other hand, with the availability of large common datasets, as the ones used in public challenges focused on medical images, more structured and comparable researches can be done. With the abundance of these data and the advent of the Vision Transformer (ViT) architecture in late 2020, a new trend in the field of medical image segmentation has spread quickly: the proposed algorithms have become more complex both in term of capacity (estimated in number of parameters), sometimes neglecting simpler, yet promising solutions.

Being conscious of my hardware limitations, the impracticability of training very large models and even the difficulties of making continuous training sessions, I had to devise a clever way to train a "good enough model" in the most effective and reproducible way, while tracking the training processes in a fashion s.t. results of different experiments were easily comparable. For this reason I leveraged deep learning frameworks for professional AI researchers for code organization and accelerating research in Medical Imaging, with the intent to create a personal baseline to confront with, and flexibly expand my solution with new parts. Many official guides, as well as research papers present long training sessions (from hundreds to thousands epochs) and the few that show how the training process evolved, share the fact that the learning curves are very noisy for most of the time and only at the very end they reach a convergence point. The only exception seems to be the ViT based architectures, but they incur more easily into overfitting. What I tried to obtain

was a flexible, relatively small model capable of learning without overfitting even with little data and in few epochs, so to have as quickly as possible a considerable amount of different "prototypes" to refine.

The work is organized as follow: ...

Chapter 2

The Challenge

The International Multimodal Brain Tumor Segmentation (BraTS) challenge is an event held in conjunction with the Medical Image Computing for Computer Assisted Intervention (MICCAI) conference. In it, prominent computational scientist and clinical researchers present their work on glioma, sclerosis and other brain injuries from every kind of point of view, but with a major focus on segmentation, prognosis and other applications for the clinical context.

Every year (from 2012), BraTS make publicy available a manually annotated dataset of preoperative brain tumor scans from different international institutions to assess the advances in the automated brain tumor segmentation task, using multiparametric MRI (mpMRI) scans. It is worth to be noted that from 2021, the challenge focuses also on a second task: the evaluation of classification methods to predict the MGMT promoter methylation status at preoperative baseline scans, since it has been identified as a strong and independent predictive factor of favorable survival in glioblastoma patients undergoing chemotherapy with alkylating agents.

Amongst brain tumors, glial tumors comprise 60% of the tumors. They are a common cause of mortality in both young and old people, with a substantial male dominance for the higher grades gliomas. Within the gliomas, Glioblastoma (GBM) and diffuse astrocytic glioma (WHO Grade 4 astrocytoma) are the most common and aggressive primary malignant brain tumors, comprising more than 16% of all primary brain and central nervous systems neoplasms. The typical survival range (prognosis) for these kind of tumors is about 1 year and the current standard of care treatment comprises surgery, followed by radiotherapy and chemotherapy. The use of MRI and diffusion tensor imaging (DTI) in preoperative planning, as well as ultrasound, CT scans, and MRI with direct stimulation during surgery, has allowed for multimodal neuronavigation and the integration of patient-specific anatomic and functional data. Despite these technologies, differentiating between normal brain and residual tumor continues to be a major challenge, and the use of an appropriate dye for fluorescence guidance has been found to be more effective then conventional neuronavigation-guided surgery alone.

The challenge requires to develop new method (or improve an existing one) being able to produce segmentation labels of the different glioma sub-regions considered for the evaluation: "enhancing tumor" (ET), "tumor core" (TC) and "whole tumor" (WT). ET is a region showing hyper-intensity in T1Gd w.r.t. healthy white matter,

but also T1 modality; TC is the bulk entailing the ET and the necrotic (NCR) parts of the tumor, and it is the section that is typically resected: surgical removal of as much of a tumor as possible (tumor debulking) has the double benefit of enhancing the effects of chemo and radiation therapies and alleviating the pain due to the symptoms. The WT covers the entire extent of the tumor, entailing TC and the accumulation of fluid in the intracellular or extracellular spaces (cerebral edema, ED).

Chapter 3

Fundamentals of magnetic resonance imaging

3.1 Medical Imaging

Medical Imaging (MI) is a process enabling the visual representation of the interior of a body for clinical analysis and medical intervention, establishes a common ground of normal anatomy and physiology to make it possible to identify abnormalities.

3.2 Nuclear Magnetic Resonance

Before deep dive into my work, let's have a small digression about the physics behind MRI in order to have an intuition on how MRI images are generated, and understand the the type of data I have worked on. Differently from computed tomography (TC) and PET scans, MRI does not use neither X-rays not ionizing radiation: it applies nuclear magnetic resonance (NMR), producing better contrast in images of soft-tissue, like the tissue in the brain.

The core idea behind NMR is to use powerful magnets to polarize and excite hydrogen nuclei of water molecules in human tissue, producing a detectable signal which result in images of the body. The specific physical characteristic of tissue visible in the image depends on how the magnetic field is being changed during the acquisition process, which consists of many repeated cycles. During these cycles the tissue magnetization is forced through a series of changes and its level that is present at the end of each cycle determines the intensity of the radio frequency (RF) signal produced and the resulting tissue brightness in the image. MR images are identified with specific tissue characteristic or blood conditions that are the predominant source of contrast; among the sources of contrast, "Magnetic characteristics of tissues" is the most common category and comprises the measurement of density of protons and the recovery process from tissue magnetization (see **Relaxation**). Images can be created in which either one of these characteristics is the predominant source of contrast: it is not usually possible to create images in which one of the tissue is the only pure source of contrast.

3.3 Radio Frequency Signal Intensity

The MRI process uses RF signals to transmit the image from the patient's body. The RF energy used is a form of non-ionizing radiation. The RF pulses are absorbed by the tissue and converted to heat and a small amount of energy is used to produce an image. The visual information that an MRI scan conveys clearly is the RF signal intensity emitted by the tissue. Bright areas correspond to tissues that emit high signal intensity and dark voids correspond to non-responding tissues. Between these two extremes resides a range of shades of gray showing contrast and differences among the tissues: when we look at an MR image, we are seeing a display of magnetized tissue, and it is the difference rates of change of magnetization level that produces much of the useful contrast at a specific "picture snapping time".

3.4 The MR Image

3.4.1 Spatial Characteristics

The MRI acquisition protocol allow the tuning of a various set of parameters to produce the appropriate spatial characteristics required by a specific clinical procedure. These characteristics include the number of slices, slice orientation and the structure within each individual slice.

Slices

A typical scan consists of a set of contiguous slices acquired simultaneously, the capacity of the set is limited by certain imaging factors and the amount of time spent for the acquisition process. The slices can be oriented in virtually any plane, but a set of oriented slices requires its own dedicated acquisition session. However, there is the possibility of acquiring 3D data from a large volume of tissue and then reconstructing slices in the different planes.

Voxels

In 3D computer graphics, a voxel represents a value on a regular grid in three-dimensional space. Each slice of tissue is subdivided into rows and columns of individual volume elements (voxel), whose size has a significant effect on image quality. Each voxel is an independent source of RF signals: this is why voxel size is a major consideration in each image acquisition.

Image Pixels

The image is also divided into rows and columns of picture elements (pixels). A pixel represent a corresponding voxel of tissue within the slice and its brightness is represent the intensity of the RF signal emitted by the tissue voxel.

3.4.2 Image Characteristics

Not all types of clinical procedure require images with the same characteristics, but the MRI system is powerful enough to allow for tremendous control over the specifics required and the overall image quality produced.

Contrast Sensitivity

Contrast sensitivity is the ability of an imaging process to produce an image of objects or tissues in the body that have relatively small physical differences or inherent contrast. The advantage of MRI w.r.t. other imaging techniques is that it has a high contrast sensitivity for visualizing differences among the tissues in the body because there are several sources of contrast: if a certain medical condition does not produce a visible change in one characteristic, there is the possibility that it will be visible in others.

Detail, Noise and Artifacts

The visibility of anatomical detail (spatial resolution) is limited by the blurring effect that occurs during the acquisition process. All medical imaging processes are affected by the blurring to some extent, but in MRI this is more accentuated. In addition to blurring, visual noise is a major issue in MRI: its presence limits the visibility of low contrast objects and differences among tissues. Most of the noise is the result of a form of random unwanted RF energy picked up from the patient's body, and the attempts to mitigate it involve necessarily some compromises with other characteristics. Artifacts are unwanted objects resulting in the processed image which do not represent an anatomical structure. They are produced by interactions or functions (such as movements) of the patient's body during the acquisition phase.

3.5 The Magnetic Field

The heart of the MRI system is a magnet that produces a strong, homogeneous magnetic field. The patent's body is placed inside the field during the acquisition procedure and it subjected to two distinct effects that combined create the image: tissue magnetization and tissue resonance.

3.5.1 Tissue Magnetization and Resonance

When immersed in a magnetic field, the tissue becomes temporarily magnetized by the alignment of the protons. The ability of MRI to distinguish between different types of tissue is based on the fact that different tissues, both normal and pathologic, will become magnetized to different levels or relax at different rates. The magnetic field also causes certain nuclei in the tissue to resonate in the RF range. The tissue then serves as both a radio receiver and transmitter during the imaging process.

3.5.2 Gradients

When the MRI system is in a resting state, the magnetic field is quite homogeneous over the region of the patient's body. To acquire data, the field must be distorted with gradients, *i.e.* changes in field strength from one point to another in the body. These changes are produced by a set of coils that are turned on and off accordingly to the acquisition procedure. In a MRI system there are typically three sets of gradient coils oriented so that gradients can be produced in the three orthogonal directions (x, y, z). Also, two or more coils can be used to produce a gradient in any direction. The three basic designs of coil are body, head and surface coils: surface one is used to receive signals from a small anatomical region to produce better image quality than is possible with the other designs.

Magnetic Direction

There are two principle directions that tissue is magnetized during the imaging process. Longitudinal magnetization is then the tissue is magnetized in a direction parallel to the direction of field. Transverse magnetization is when the direction of tissue magnetization is at a 90° angle w.r.t. the direction of the magnetic field and is in the transverse plane. Note that the actual direction of magnetization is not limited to longitudinal or transverse: it can have both components and have distinct characteristics that must be considered independently.

Magnetic Flipping

When a 90° pulse is applied to the longitudinal magnetization, it is reduced to zero (*saturation*) and flipped into the transverse plane, producing transverse magnetization: an *excited* condition (see **Excitation**).

Longitudinal Magnetization And Relaxation

When the magnetization is redirected by an RF pulse, it will recover its original orientation over a period of time (*relaxation time*). The convention is to specify the relaxation time in terms of the time required for the magnetization to reach the 63% of its maximum (value used for mathematical considerations and practical purposes). This time, the *longitudinal relaxation time* is named T1. Each tissue has its own T1 and the relevant aspect is that tissues with short T1 will have the highest level of magnetization (*i.e.* will be the brightest) in T1-weighted images when the picture is snapped during the relaxation period.

Transverse Magnetization And Relaxation

After the conversion to transversal magnetization, the excited condition quickly decay accordingly to a specific relaxation time named T2. Different tissues have different T2 values and the level of magnetization at snapping time will be used again as a source of contrast in the MR image. Transverse magnetization aside from contrast generation has also the purpose of generating the RF signals emitted by the tissue: each imaging cycle must conclude with transverse magnetization to produce the

RF signal to form the image. Since Transverse Magnetization relaxation is actually a decay, T2 is the time required for 63% of the initial magnetization to dissipate. In general, this is the reason for a T2-weighted image appear to be a reversal of its T1-weighted version. The two factors that contribute to the dephasing of the nuclei, hence to the T2 relaxation, are the *spin-spin interactions* among the nuclei, and the inhomogeneity of the magnetic field. As a consequence, the true relaxation characteristics of a tissue are masked and the real relaxation time is named T2*, resulting much less then T2, and the transverse magnetization disappears before T2 contrast can be formed and sensed. The spin echo process is used to compensate the rapid relaxation time: a 180° pulse is applied to the tissue and the protons are flipped around an axis in the transverse plane; this cause an inversion of direction of rotation and make the fast spinning protons realigning with the slower ones, producing again a transverse magnetization (*echo event*) that builds up to a level dictated by T2 characteristics of the tissue.

Inversion Recovery and FLAIR MR imaging

Following the same principle of the spin echo process applied for the acquisition of T2 contrast, inversion recovery (IR) is a spin echo method used for specific purposes, such as obtain a high level of T1 contrast, suppress the signals of fat and fluids. This method applies an additional 180° pulse to the conventional spin echo sequence, inverting the direction of the longitudinal magnetization. The recovery of the magnetization level starts from a negative value and an additional time interval is associated with the inversion recovery pulse sequence: it is the time between the initial 180° pulse and the 90° pulse that zeros the longitudinal magnetization. This time is named Time after Inversion (TI) and can be used as a contrast control tool. The fluid-attenuated inversion recovery (FLAI) is an IR process in the spin echo methods family to null fluids in the image acquisition: it can be used in brain imaging to suppress cerebrospinal fluid (CSF). On T2 weighted images, astrocytoma and CSF have extensive areas of fairly homogeneous high signal. On T2-FLAIR, instead, the majority of these areas become relatively hypointense in signal due to incomplete suppression and, at the margins of the tumour, a rim of hyperintensity is usually seen.

3.5.3 Nuclear Magnetic Interactions

A magnetic nucleus is characterized by a magnetic moment. In the absence of a strong magnetic field, the moments are randomly oriented in space. When the magnetic nuclei are placed in a magnetic field, they experience a torque that encourages them to align with the direction of the field. In a human body, the number of nuclei capable of this alignment is proportional to the field strength. Once the nucleus is aligned, its nuclear magnetic moment precesses about the axis of the magnetic field. It is this wobbling motion that makes a nucleus sensitive and receptive to incoming RF energy when the RF frequency matches the precession rate (see **Larmor Frequency**).

Excitation

In MRI a RF pulse is used that flips some of the nuclei alignment away from the direction of magnetic field, into its transverse plane. This motion places the nuclei in an *excited* state. In this state, the precession is transformed into a motion around the axis of the field and produces the RF signal that is collected to form the MR image.

Relaxation

As already mentioned, to acquire a full MR image a patient must undergo several cycles of MR. During these cycles, the tissue magnetization is flipped into an unstable condition and then it is allowed to recover: this recovery process is known as *relaxation*. The time required to "relax" depends on the physical characteristic of the tissue and can then be used to distinguish among normal and pathological tissues. Each tissue is characterized by two relaxation times: T1 and T2 that may be chosen to be the predominant source of contrast in a particular scan, which is named after the chosen time.

Larmor Frequency

The resonant frequency (*Larmor frequency*) of a nucleus is determined by a combination of nuclear characteristics and the strength of the magnetic field. The fact that different nuclides have different resonant frequencies means that most procedures can "tune in" with only one chemical element at a time.

Chapter 4

Related Works

Chapter 5

SOTA Architectures and Procedures

Chapter 6

The Data

Chapter 7

Experimental Design and Result

Chapter 8

Conclusions and Future Work