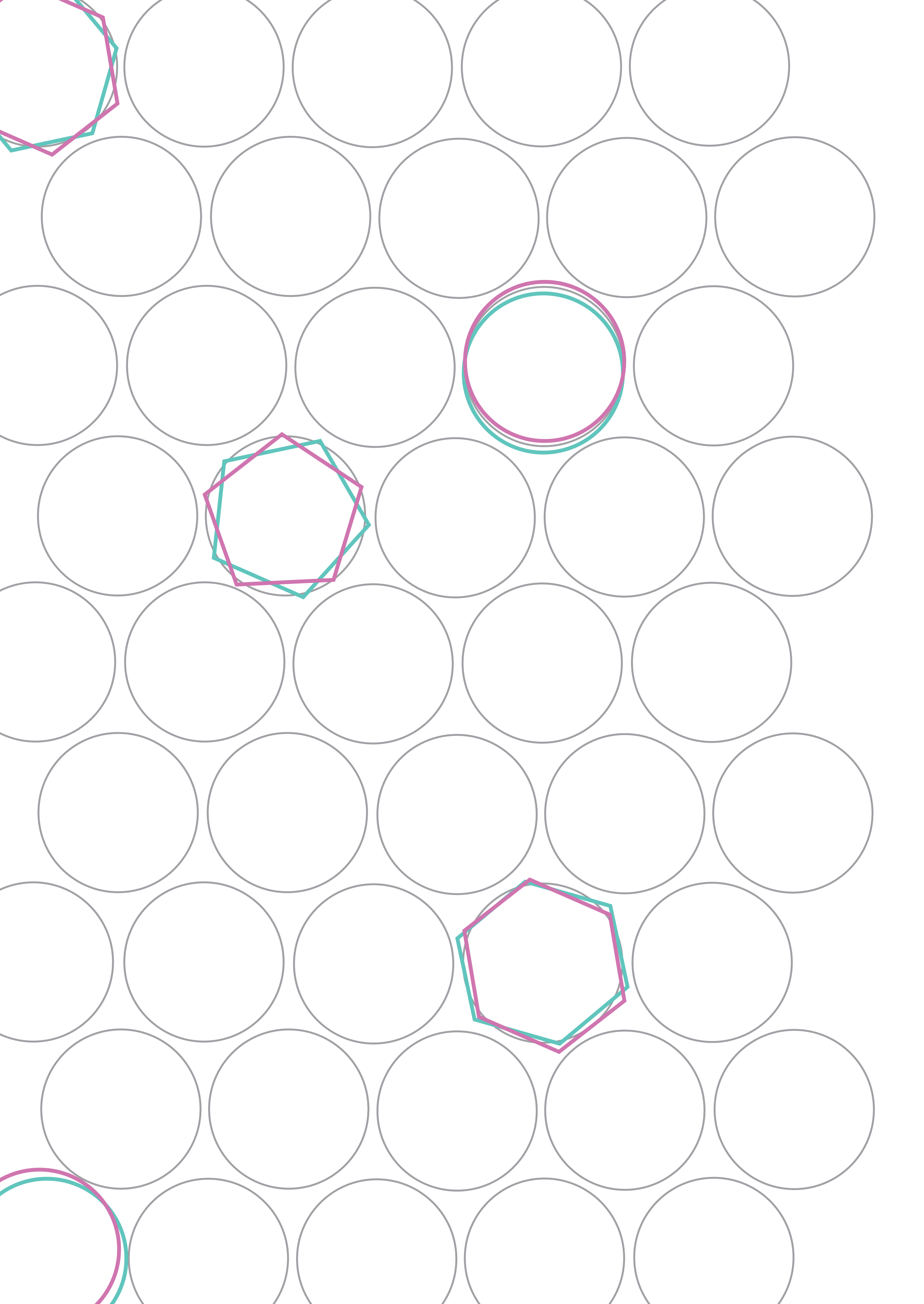

Chapter 1

Introduction

TUMOR SEGMENTATION IN RADIOTHERAPY

Cancer is the second leading cause of death worldwide, accountable for nearly 10 million deaths in 2020 [1]. Radiotherapy (RT) is one of the main modalities of cancer treatment, being indicated as part of the treatment for approximately half of the patients [2–4]. RT consists of the use of high doses of ionizing radiation to damage the DNA of the cancer cells, leading to their death. Simultaneously, it is intended that the least amount of radiation possible is given to the surrounding healthy tissue. Therefore, precise localization of the tumor and of the surrounding organs is critical for the RT treatment planning.

To visualize a patient's anatomy, an image of the patient is acquired using common imaging modalities such as computed tomography (CT), magnetic resonance (MRI), or positron emission tomography (PET). On these images, the pixels corresponding to the relevant structures are identified and labeled. This process is known as segmentation, delineation or contouring. The structures to segment are the target of the RT treatment and the organs to spare (known as organs at risk - OARs). Once these segmentations are made, the dose intended to each region of the patient anatomy can be calculated. Given that the segmentation step is one of the first steps of the RT workflow, it can influence the quality of the entire treatment [5].

The primary target of the RT treatment is typically the visible extent of the tumor on the image of the patient, known as gross tumor volume or GTV. In addition, to address any hidden microscopic disease that is not visible on the imaging, an extra margin is defined, known as clinical target volume or CTV. Furthermore, a safety margin is often defined to account for any geometrical uncertainties that may arise from the whole treatment pipeline, the planning target volume or PTV. In certain tumor sites, such as cervix or rectum, the CTV is defined following anatomical structures rather than as a margin around the tumor. Arguably, both the tumor or targets that are not defined as anatomical structures can be considered harder to segment than those anatomical structures, because their locations, shapes and sizes tend to be more variable.

In the current clinical practice, the segmentation of the tumor is carried out manually by physicians, which presents two issues. The first issue is that manual segmentation of the tumors is a time-consuming process, often regarded as one of the most time-intensive steps in the RT workflow [6] or even referred to as a bottleneck of the entire pipeline [7]. This is partly due to the complexity of the tumor segmentation process: it involves the physician scrolling through the 3D image of the patient and annotating the points that correspond to the tumor in a slice-by-slice manner. Additionally, physicians often integrate information from various sources, including clinical reports, other images, and clinical examinations of the patient.

Nowadays, it is aimed to deliver the RT treatment more accurately by considering anatomical changes that happen between RT planning and delivery or even during the course of the treatment [8], thereby potentially improving local control and reduce treatment-related toxicities. This approach is referred to as 'adaptive RT' or ART and it can be carried out in two ways: offline or online. Online ART (OART) specifically addresses changes that occur on the day of treatment [9]. A key step of OART is the daily (re-)segmentation of the relevant structures for the treatment plan, a process in which both the patient and the entire treatment staff are waiting while these segmentations are performed. Given the increased interest in OART, the clinical burden of segmentation is expected to increase even more in the coming years.

The second issue with manual segmentation of the tumors is that it depends on the observer. Inter-observer variability has been reported in a myriad of tumor sites, such as head and neck [10,11], esophagus [12], lung [13,14], rectum [15], or cervix [16]. Potential sources of this variability include, but are not limited to, the lack of clear boundaries of the tumor, the lack of standardized delineation guidelines or differences in experience among different observers. These deviations on the segmentations of the targets can affect the RT outcome in terms of RT-related toxicity, overall survival and local recurrence of cancer [17].

DEEP LEARNING FOR AUTOMATIC SEGMENTATION OF TUMORS IN RADIOTHERAPY

Deep learning techniques, specifically convolutional neural networks (CNNs), are currently considered state of the art in computer vision tasks, including segmentation. Therefore, CNNs have been the preferred method in the recent years for automatically segmenting the structures required for RT treatment, namely the targets and the OARs. For the case of OARs, CNNs have already been shown to outperform previous methods of automatic segmentation for several anatomical sites [15–18]. Furthermore, they are already incorporated in the software released by several companies [22–24].

For the case of the targets, commercial solutions are only available to segment the CTV in a few cases, when it is defined as an anatomical structure, like the prostate, the head and neck lymph nodes or the vessels of the pelvis [24,25]. Automatic segmentation of the tumors is, however, still uncommon and their clinical adoption is non-existent. Promising results have been reported [26,27], often achieving arguably high performance (Dice Scores between 0.73-0.94), but the segmentations resulting from these methods still can be unacceptable in some of the cases. Given that the tumors receive the highest dose of radiation during treatment, errors in their automatic segmentation could result in unacceptable treatment plans for clinical practice [17].

Typically, the set up for automatic segmentation of tumors is done in a fully supervised manner [26,28]. During a training phase, the CNN is shown images of different patients (i.e. input images) and asked to provide a segmentation of the tumor. In each training iteration, the proposed automatic segmentation is compared to the manual segmentation made by the doctors, and the error is quantified. The function that quantifies this error is known as loss or error function. This error function is used to modify the parameters of the CNN, until the proposed segmentations are similar to the manual segmentations. Important lines of research in the field of automatic segmentation of medical structures focus on finding the best input images to the network to achieve the most accurate segmentations [29] and determining the optimal loss functions to train the CNNs [30] .

Another line of research related to the automatic segmentation of tumors is the automation of the quality assurance of the automatic segmentations. Because these segmentations are not yet perfect [26,28,31], they still need to be checked by physicians, which hinders the promise of automatic segmentations for time saving. Recently, several works have focused on developing uncertainty maps, that show where the CNN is uncertain about its predictions [32–35]. These uncertainty maps have been shown to correlate with the areas where the network failed [36,37], and could thereby be used to flag the segmentations that would require a check.

AUTOMATIC SEGMENTATION OF TUMORS ON MRI

Due to its superior soft tissue contrast to other image modalities, MRI is a desirable image modality for RT treatment planning purposes. Moreover, each MRI sequence enables visualization of different tissue types, even with the possibility of imaging biological and functional processes through quantitative MRI techniques. This level of versatility renders MRI a valuable imaging modality and distinguishes it from other imaging techniques like CT. Furthermore, the acquisition of MRI does not involve the exposure of the patient to ionizing energy, as opposed to CT or PET. In recent years, systems that combine a MRI scanner and a RT linear accelerator for RT treatment delivery have been made commercially available [38], such as the Elekta Unity (Elekta AB, Stockholm, Sweden) MR-Linac system or the ViewRay MRIdian system (Viewray Inc., Oakwood, OH). Thus, there is an increased interest in MRI-only pipelines for RT treatment planning.

Partly due to its relative novelty in the field of RT as main imaging modality, the research on automatic segmentation of tumors on MRI is even less spread than on CT or PET-CT [26,28]. The lack of large MRI-based datasets, which are needed to train the data-hungry deep learning methods, can also explain the scarcity of literature on the topic. One notable exception is the case of brain tumors, where MRI is the dominant image modality [39]. Two tumor sites that showcase the potential of MRI as main imaging modality for automatic tumor segmentation are the oropharyngeal cancer, due to the high prevalence of soft tissue in the area; and the cervical cancer, for which MRI-guided radiotherapy is the standard treatment option [40].

One of the main treatment modalities for oropharyngeal cancers is RT, being indicated in more than 70% of the patients [41]. A challenge in RT for this tumor site is the proximity of many OARs, which makes the accurate segmentation of the tumor all the more critical to avoid potential damage to those organs. Furthermore, high interobserver variability has been reported for this tumor site, reaching differences in the volumes delineated by the different observers of up to a factor of 10 [11]. Automatic segmentation of the tumors is therefore a desirable solution for this tumor site [42], not only to decrease the clinical burden when treating this tumor but also to provide more consistent segmentations among different physicians or even hospitals.

Locally advanced cervical cancer is typically treated with a combination of external beam radiotherapy (EBRT), concomitant chemotherapy, and 3 to 4 fractions of brachytherapy (BT) [43]. BT consists of the placement of radioactive sources directly into or next to the tumor. For the case of cervical cancer, this is achieved with an applicator which is inserted into the patient before the BT treatment. Afterwards, the MRI images are acquired, and the patient must remain immobilized in bed while the necessary structures (such as target volumes and organs at risk) are manually segmented, and a treatment plan is devised. This is highly uncomfortable for the patient and logistically complex, which makes automatic segmentation even more critical than for other RT treatment modalities.

AIM AND THESIS OUTLINE

Despite the versatility of MRI in visualizing anatomical structures and tumor tissue, there is a lack of research on automatic segmentation of tumors on MRI. Current deep learning methods still fail to produce accurate tumor segmentations in certain cases, particularly on MRI. Physicians would still need to manually review and potentially correct the automatic segmentations, which limits the promised time-savings and clinical applicability of these methods. Given the complexity of the task, we hypothesized that training a CNN with a considerably large dataset would achieve clinically acceptable automatic segmentations. However, in the medical field, gathering extensive datasets is challenging due to the expensive and logistically complex data acquisition process, as well as the need for highly skilled professionals to manually label the data. The goal of this thesis was to develop automatic segmentation techniques for tumors in MRI images that deliver clinically acceptable segmentations using clinical MRI datasets. Additionally, we aimed to automate the quality assurance of the automatic segmentations, thereby maintaining the time-saving benefits even in cases where the network fails. The different automatic segmentation methods were applied in two different tasks: the automatic segmentation of the oropharyngeal primary tumor in multiparametric diagnostic MRI images (chapters 2 and 3) and the automatic segmentation of the cervical cancer gross tumor volume (GTV) in the MRI images of the BT treatment images (chapters 4 and 5).

In the oropharyngeal cancer segmentation task, we investigated three different strategies to introduce prior information in the neural network training design. Firstly, we studied the effect of using different anatomical MRI sequences as input to the network (Chapter 2), similarly to how radiologists use the different sequences to delineate the tumor themselves. Secondly, we investigated the use of different loss functions (Chapter 3). Finally, we studied the effect of reducing the context around the tumor, first by proposing a semi-automatic approach in which human observers approximately located the tumor and a CNN segmented the tumor (Chapter 2) and then fully automatizing it into a two-stage approach, that split the task in detection and segmentation (Chapter 3).

In the cervical cancer segmentation task, we validated the quality of the segmentations obtained with a state-of-the-art segmentation framework not only geometrically but also with dose-volume parameters and investigated whether there were differences for clinically relevant parameters, such as volume or tumor stage (Chapter 4). Furthermore, even though the proposed method performed adequately on average, the network still failed in some cases. We identified a metric to predict the quality of the automatic segmentations (Chapter 5). Such a metric can potentially flag segmentations that would require a manual check and could therefore help with the clinical implementation of the automatic segmentation methods.