Closing the loop: fully automated radiotherapy planning for head and neck cancer

Gao, R.

1
Introduction

This thesis focuses on advanced techniques in radiotherapy, with particular emphasis on generalizable medical image segmentation models, deep learning–based dose prediction, and automated treatment plan generation through dose mimicking. This chapter introduces the clinical and methodological background relevant to this thesis, including:

- The clinical principles, workflow, and technological evolution of external beam radiotherapy, with emphasis on photon-based techniques such as VMAT;
- The anatomical and planning challenges associated with head and neck radiotherapy;
- The formulation of radiotherapy treatment planning as a multi-criteria optimization problem;
- Recent advances in automated radiotherapy planning, including deep learning–based dose prediction and the generation of clinically deliverable treatment plans;
- An overview of the thesis structure and the main contributions presented in the subsequent chapters.

1.1 External beam radiotherapy

Cancer treatment typically involves multiple approaches, including surgery, systemic therapy (e.g., chemotherapy and immunotherapy), and radiotherapy (RT). Among these, radiotherapy plays an important role in modern oncology, with approximately half of all cancer patients receiving RT at some point during the course of their disease [1, 2]. Unlike surgery, which removes tumor tissue directly, or systemic therapies that affect the whole body, radiotherapy is a local treatment that delivers ionizing radiation to carefully defined target areas. The goal is to cause irreversible DNA damage in tumor cells while minimizing dose to surrounding healthy tissues [3]. In clinical practice, radiotherapy is mainly delivered in two ways: external beam radiotherapy, where radiation is delivered from outside the patient, and brachytherapy, where a radiation source is placed inside or close to the tumor. This thesis focuses on external beam radiotherapy.

A typical external beam radiotherapy workflow consists of several sequential steps, including patient intake and consultation, planning CT acquisition, anatomical delineation of target volumes and organs-at-risk (OARs), treatment planning, irradiation, and post-treatment follow-up [4]. These steps are closely interconnected: the planning CT and delineated structures provide the anatomical basis for treatment planning, the resulting plan must be delivered accurately during irradiation, and follow-up is needed to assess tumor response and treatment-related toxicity. This thesis mainly focuses on automating the delineation and treatment planning components within this workflow.

External beam radiotherapy primarily relies on photon or proton beams. Photon beams, generated by medical linear accelerators, deposit dose gradually along their path with both entrance and exit dose. Proton beams, in contrast, exhibit a Bragg peak, allowing dose deposition to be concentrated at a defined depth with minimal exit dose [5]. While proton therapy offers potential dosimetric advantages, photon therapy remains the most widely available and clinically used modality worldwide [6]. This thesis focuses exclusively on photon radiotherapy.

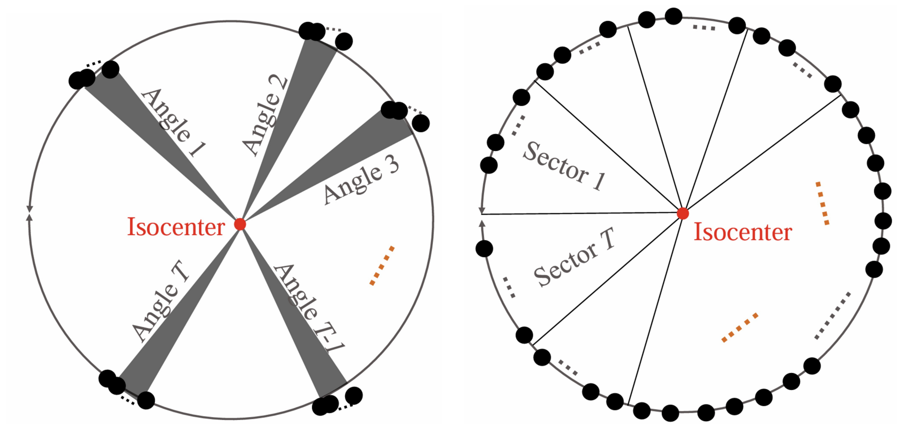
Technological advances have significantly improved photon delivery. Intensity-Modulated Radiation Therapy (IMRT) introduced fluence modulation through multi-leaf collimators (MLCs), enabling highly conformal dose distributions. Volumetric Modulated Arc Therapy (VMAT) further extended this concept by delivering radiation continuously during gantry rotation, dynamically modulating MLC positions, dose rate, and gantry speed, as shown in Figure 1.1. Compared to static-field IMRT, VMAT typically achieves similar or improved plan quality while reducing treatment delivery time [7].

These technological developments have substantially increased plan quality, but they have also increased planning complexity. In this thesis, we mainly focus on VMAT, which represents current clinical practice.

1.2 Challenges of head and neck radiotherapy

Head and neck cancer (HNC) represents one of the most challenging disease sites in radiotherapy. This region contains many OARs densely distributed around tumors that often exhibit complex and irregular geometries [9]. Accurate anatomical delineation is therefore essential for treatment planning, as segmentation of planning target volumes (PTVs) and OARs defines the geometric basis for subsequent dose optimization.

However, segmentation in the head and neck region remains challenging. Many critical structures are small, thin, or poorly visible on CT, the primary imaging modality used in the radiotherapy workflow. Multimodal imaging such as MRI or PET is frequently incorporated to improve visualization, but this also introduces additional



(a) IMRT (b) VMAT

Figure 1.1: Comparison of IMRT (a) and VMAT (b). IMRT plans consist of multiple fixed beams, each containing several control points at identical gantry angles. In contrast, VMAT delivers radiation continuously along a designated arc, where control points are evenly distributed and may be grouped into multiple sectors during planning and optimization. Figure adapted from [8].

complexity in image registration and cross-modality interpretation. Substantial inter-observer variability has therefore been reported in clinical contouring practice [10]. These challenges have motivated the development of robust and generalizable segmentation approaches that can operate reliably across imaging modalities and anatomical variations, providing a consistent foundation for downstream automated workflows. Recent advances in general-purpose segmentation models, such as the Segment Anything Model (SAM) [11], have demonstrated strong generalization across domains, although adaptation to medical imaging remains an active area of research [12, 13, 14, 15, 16].

Even with accurate segmentation, treatment planning in this region remains intrinsically difficult. Effective treatment requires sharp dose falloff between the PTVs and nearby OARs, as even small dose deviations may lead to clinically significant side effects, including xerostomia, dysphagia, or neurological injury. For example, irradiation of the salivary glands may reduce saliva production and cause chronic dry mouth [17]. The presence of multiple competing dose constraints and the need for rapid dose falloff further increase the complexity of treatment planning for HNC.

1.3 Treatment planning as a multi-criteria optimization problem

Radiotherapy treatment planning is formulated as an inverse problem. Given a patient's anatomy and prescribed dose, the planner must determine machine parameters

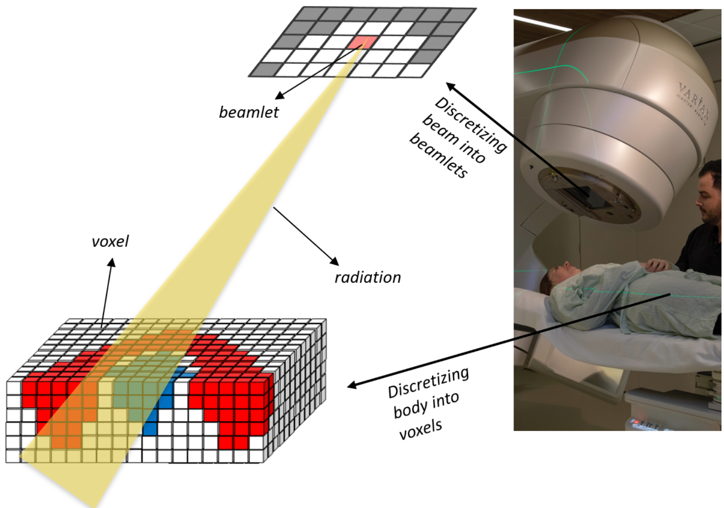


Figure 1.2: Mathematical modeling of radiotherapy dose calculation, with the beam decomposed into beamlets and the patient represented by voxels. Figure reproduced from https://github.com/PortPy-Project/PortPy.

that produce an optimal dose distribution. Mathematically, dose calculation can be expressed as $d = Ax$, where d represents voxel dose, x denotes beamlet intensities, and A is the dose influence matrix describing the physical relationship between beamlets and voxel dose deposition [18, 19] (Figure 1.2). This discretized beamlet–voxel model forms the computational foundation of radiotherapy planning algorithms.

Because treatment goals involve competing clinical priorities, radiotherapy planning is inherently a multi-criteria optimization (MCO) problem [18, 19]. Typical objectives include maximizing tumor target coverage while minimizing dose to multiple OARs, which are often conflicting. The set of optimal trade-off solutions is commonly described by the Pareto surface [18], where no objective can be further improved without degrading at least one other objective. Clinical plan evaluation frequently relies on dose-volume histogram (DVH) metrics that reflect these trade-offs between target coverage and organ sparing.

In practice, planners iteratively adjust objective weights and constraints to navigate the Pareto surface of achievable solutions. As delivery techniques such as VMAT introduce a larger number of decision variables and delivery constraints, the search space becomes even more complex.

Furthermore, an optimal dose distribution must ultimately be deliverable, meaning it can be translated into physically feasible MLC leaf trajectories and machine parameters for VMAT delivery. Therefore, automated planning must address not only optimality in dose space, but also feasibility in machine parameter space [8, 19].

1.4 Automated radiotherapy planning

The increasing complexity of modern radiotherapy planning has led to growing interest in automated solutions. Advanced delivery techniques such as VMAT involve high-dimensional parameter spaces, and treatment planning often requires iterative adjustment of optimization objectives to balance target coverage and OAR sparing. This trial-and-error process is time-consuming, cognitively demanding, and dependent on planner experience, which may lead to variability in plan quality [20, 21]. Automated approaches therefore have the potential to improve efficiency, consistency, and accessibility of high-quality radiotherapy planning.

Recent advances in deep learning have enabled three-dimensional dose prediction directly from patient anatomy [22, 23, 24, 25, 26, 27, 28]. Rather than manually navigating the multi-criteria optimization process, neural networks can learn the mapping from anatomical structures to clinically acceptable dose distributions using historical treatment plans. Such models can implicitly capture trade-offs between target coverage and normal tissue sparing, offering a potential pathway toward efficient and consistent automated planning.

However, predicted dose distributions exist only in voxel space and are not directly deliverable. Practical VMAT delivery requires explicit machine parameters, including multi-leaf collimator trajectories, gantry motion, and dose rate modulation, all subject to mechanical and physical constraints. Bridging the gap between predicted dose distributions and physically deliverable treatment plans is therefore essential for clinical implementation.

One practical strategy is dose mimicking [29, 30], in which the predicted dose distribution is incorporated into a treatment planning system as a reference objective to guide inverse planning and generate deliverable machine parameters. By combining dose prediction with dose mimicking, it becomes possible to move from patient anatomy to clinically deliverable VMAT treatment plans within an automated planning framework.

1.5 Thesis outline

This thesis investigates automated radiotherapy treatment planning for HNC, with a primary focus on deep learning–based dose prediction and the generation of clinically deliverable VMAT plans through dose mimicking. In this thesis, automated radiotherapy planning mainly refers to the stages from dose prediction to deliverable plan generation.

In addition to the radiotherapy planning components, this thesis also explores generalizable medical image segmentation methods. Although the segmentation study presented in Chapter 2 is developed for broader medical imaging scenarios, robust and scalable segmentation models are highly relevant to radiotherapy workflows,

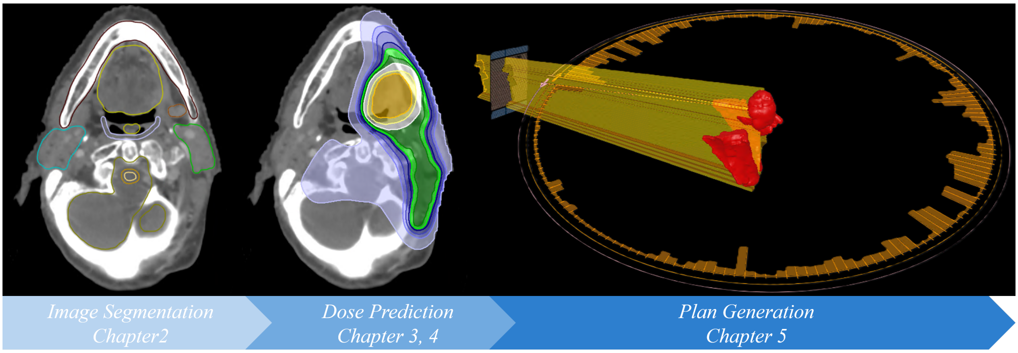


Figure 1.3: Overview of research topics in this thesis.

where accurate anatomical delineation forms the basis for reliable treatment planning and automation. Figure 1.3 provides an overview of the research topics and their relationships.

The individual contributions are structured as follows:

- **Chapter 2** presents MCP-MedSAM, a lightweight modality-agnostic medical image segmentation foundation model. The chapter addresses multi-modality medical segmentation challenges and establishes a scalable delineation framework applicable across anatomical sites and imaging modalities, including but not limited to radiotherapy.
- **Chapter 3** investigates key factors influencing deep learning–based 3D dose prediction for head and neck photon therapy, including input and dose-grid resolution, input representation, loss functions, model architecture, and the impact of noise on model performance.
- **Chapter 4** proposes a clinical DVH metric loss (CDM loss) that jointly incorporates differentiable formulations of common *D–metrics* and differentiable surrogates of *V–metrics*. Additionally, a lossless bit-mask encoding scheme is introduced to efficiently represent a large number of overlapping regions of interest, significantly improving training efficiency.
- **Chapter 5** integrates the trained dose prediction model into a commercial treatment planning system (RayStation) and performs automated dose mimicking to generate clinically deliverable VMAT plans.
- **Chapters 6 and 7** summarize the overall contributions, implications, and future research directions in English and Dutch, respectively.