

MENG INDIVIDUAL PROJECT

DEPARTMENT OF COMPUTING

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

**Transfer Learning for bespoke automatic
contouring of cervical cancer radiotherapy
planning**

Author:
Anton Zhitomirsky

Supervisor:
Ben Glocker

Second Marker:
TODO

January 10, 2024

Abstract

Clinicians target cancerous tumours by studying 3D contrasting images of cancerous tumours and surrounding soft tissues to plan targets for radiation therapy. The Royal Marsden Hospital is a key contributor of data for this project, which uses this approach to delineate tumours for cervical cancers. Typically after a gross tumour volume (GTV) is extrapolated from the relevant imaging modality, clinicians append tailored safety margins to also account for the microscopic cancerous spreads not visible in the scan to generate the planned target volume (PTV).

The PTV area has to be generous enough to attempt to treat the problem in one-shot, yet conservative enough to not harm surrounding healthy tissue with radiation over the course of the treatment. Compounded with small sample size of labelled data this proposes a significant challenge for developing deep-learning segmentation models to identify an optimal PTV.

Thus we propose a transfer learning strategy to utilize imaging models in similar domains to attempt to learn from the limited input size to provide clinicians with a faster and more accurate segmentation method.

Contents

1	Introduction	1
1.1	Clinical Context	1
1.2	Motivation	1
1.3	Current Solutions	1
1.4	Outline of Report	1
2	Background	2
2.1	Clinical Context	2
2.2	Vanilla Image Segmentation Models	2
2.2.1	Convolutional Neural Networks (CNN)	2
2.2.2	U-Net	2
2.2.3	nnU-Net	2
2.2.4	Traditional Limitations	2
2.3	Transfer Learning	2
2.3.1	Intuition	2
2.3.2	Total Segmentator	2
2.3.3	UniverSeg	2
2.3.4	SAM	2
2.4	Summary	2
3	Base Line and Data	3
3.1	Data	3
3.1.1	Available Data	3
3.1.2	Data Cleaning	3
3.2	Evaluation Metrics	3
3.2.1	Geometric	3
3.2.2	Honorable mentions	3
3.3	Baseline Results	3
3.3.1	nnU-Net	3
3.3.2	Total Segmentator	3
3.3.3	UniverSeg	3
3.4	Summary	3
4	Proposal	4

5 Ethics	5
5.1 Patient disclosures	5
5.2 Using the tool	5
Bibliography	7

Chapter 1

Introduction

1.1 Clinical Context

1.2 Motivation

1.3 Current Solutions

1.4 Outline of Report

Chapter 2

Background

2.1 Clinical Context

2.2 Vanilla Image Segmentation Models

2.2.1 Convolutional Neural Networks (CNN)

2.2.2 U-Net

2.2.3 nnU-Net

2.2.4 Traditional Limitations

2.3 Transfer Learning

2.3.1 Intuition

2.3.2 Total Segmentator

2.3.3 UniverSeg

2.3.4 SAM

2.4 Summary

Chapter 3

Base Line and Data

3.1 Data

3.1.1 Available Data

3.1.2 Data Cleaning

3.2 Evaluation Metrics

3.2.1 Geometric

3.2.2 Honorable mentions

3.3 Baseline Results

3.3.1 nnU-Net

3.3.2 Total Segmentator

3.3.3 UinverSeg

3.4 Summary

Chapter 4

Proposal

Chapter 5

Ethics

This project involves very intimate and personal information of many female patients. Researchers may collaborate with third-parties by providing anonymized data which may not be reverse engineered back to the patient. The lack of this effort may result in “stigma, embarrassment, and discrimination” [6] if the data is misused.

5.1 Patient disclosures

The Royal Marsden Hospital doesn’t require “explicit consent” for sharing collected clinical data with outside entities as long as the patient is made aware of the ways their “de-identified/anonymized” data may be used [7]. Formalities are also arranged with Imperial Collage’s Medical Imaging team such as acting as “ethical data stewards” [4]. Without such disclosure and anonymisation of data, patients may be reluctant to provide candid and complete disclosures of their sensitive information, even to physicians, which may prevent a full diagnosis if their data isn’t maintained in an anonymous fashion.

The MIRA team acts as responsible data stewards by storing anonymized data within a folder on the college network. All provided data was anonymized by the Royal Marsden Hospital and sent to team MIRA in the NIfTI file format which discloses no personal identifiable information, as defined by GOV website [3]. This folder contains security measures which limit the availability of data only to those with specific access rights. Furthermore, operating on the preamble of de-identified data further reduces individual patient risk in the event that data is ever brought outside the confines of this folder.

5.2 Using the tool

The applications of this tool bode well in the healthcare ecosystem as the community slowly realizes the importance of AI-powered tools for the next generation of medical technology. Radiology has been one application that has been most welcoming of the new advances in technology as there is potential for substantial aid by reducing manual labor, increasing precision and freeing up the primary care physician’s time [1].

Yet, it is too early to take result the medical tool as gospel. For current cervical radiotherapy delineation tools, only 90% of the output is considered as acceptable for clinical use [5]. The remainder therefore has the potential to cause more harm than good if not checked properly. For example, overlap of a PTV onto an organ-at-risk may invoke a cascade of negative effects for the patient. A potential cause may be the lack of multivariate analysis, where an oncologist would need to consider a variety of data, whereas this model only considers a single point of evidence (results of an imaging modality). Therefore, a degree of supervision required from physicians has to be established if this tool is to be used in practice. Oncologists will be required to reverse-engineer results of the ‘black-box’ to verify

why a decision has been made. Secondly, the responsible party for incorrect decisions made by DL tools should also be determined [2].

Bibliography

- [1] Amisha et al. “Overview of artificial intelligence in medicine”. In: *Journal of Family Medicine and Primary Care* 8.7 (2019). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6691444/>.
- [2] Zi-Hang Chen et al. “Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine”. In: *Cancer Communications* 41.11 (2021). URL: <https://onlinelibrary.wiley.com/doi/10.1002/cac2.12215>.
- [3] *Data Protection Act 2018*. URL: <https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted>.
- [4] David B. Larson et al. “Ethics of Using and Sharing Clinical Imaging Data for Artificial Intelligence: A Proposed Framework”. In: (2020). URL: <https://pubs.rsna.org/doi/full/10.1148/radiol.2020192536>.
- [5] Zhikai Liu et al. “Development and validation of a deep learning algorithm for auto-delineation of clinical target volume and organs at risk in cervical cancer radiotherapy”. In: *Radiotherapy and Oncology* 153 (2020).
- [6] Nass SJ, Levit LA, and Gostin LO. “The Value and Importance of Health Information Privacy”. In: (2009). URL: <https://www.ncbi.nlm.nih.gov/books/NBK9579/>.
- [7] The Royal Marsden NHS Foundation Trust. “Privacy Note”. In: (2023). URL: https://rm-d8-live.s3.eu-west-1.amazonaws.com/d8live.royalmarsden.nhs.uk/s3fs-public/2023-10/T22020ac_Revised%20privacy%20policy_V1_AW_WEB.pdf.