Imperial College London

MEng Individual Project

Department of Computing

Imperial College of Science, Technology and Medicine

Transfer Learning for Deep Learning Radiotherapy Planning

Supervisor:
Prof Ben Glocker

Author:
Anton Zhitomirsky

Second Marker:
Dr Thomas Heinis

June 8, 2024

Abstract

Cervical cancer remains as one of the top cancerous diseases to affect women. To treat it, Oncologists plan a contour for therapy after obtaining 3D contrasting images of soft-tissue organs at risk and tumorous areas.

Auto-segmentation differs from Auto-contouring tasks due to lacking clinical knowledge surrounding the location of the cancer and biological spreading patterns. Instead of trivially contouring visible macroscopic tumour masses on a scanned patient, a clinician requires also to adjust for microscopic spreads and finally error margin spreads. This target volume should aim to treat the disease in one-shot and not affect any organs-at-risk.

The scientific community has tried to automate this task using current architectural standards such as CNNs or U-Net based algorithms. However, no studies yet consider Transfer Learning as an approach to solving this issue. This report investigates this architectural challenge to contribute to the total pool of deep learning auto-segmentation models in a communal effort to save resources of medical institutions.

You can find the most up-to-date version of the report *here*.

Contents

Chapter 1

Introduction

1.1 Clinical Context

In 2017, Cervical Cancer accumulated 530,000 new cases annually, with 270,000 deaths, making it the fourth most common malignancy diagnosed in women worldwide [1]. A common treatment mechanism involves radiation therapy which targets cancerous cells in a clinically defined target area with beams of high energy (Section **??**). This treatment is tedious, as it is estimated that an oncologist needs 90-120 min to delineate target areas for radiotherapy [2].

Radiotherapy has become a great option due to high resolution X-ray or CT scans which produce high contrasting images of the damaged and surrounding soft tissue [3]. Physicians then use this 3D scan to plan a target volume for the radiation therapy surrounding the tumor in hopes of killing it and not damaging the surrounding tissue.

Areas are therefore constructed based on an Oncologist knowledge about the particular cancer to determine target structures, structures we need to protect (organs-at-risk), and areas where each particular cancer is likely to spread to [4]. These areas are delineated onto a patient scan and used as a radiotherapy target used for treatment.

Accurate scans have been provided to see if an AI model can learn cervical cancer CTV patterns adjacent to clinical knowledge and oncologist prior knowledge. Training models which produce substructures required for radiotherapy target volumes would overall save time and improve consistency within the radiotherapy planning process [4].

Chapter 2

Background

2.1 Current Limitations

However, our problem is not accurately solved with the methods mentioned. This is due to a handful of independent details which require more careful planning and engineering.

2.1.1 Data Size

The data quantity supplied is a limiting factor for creating a robust model. We are given 100 labeled data elements across 5 classes. Without vast collection of knowledge, it is hard for an application to create a model which generalizes well to the total population, especially in a very specific and bespoke use case as radiotherapy planning for cervical cancer.

2.1.2 Bespoke Application

Another issue lies in the bespoke nature of this application. Most pre-trained networks currently run segmentation on structures are that are more obvious in a given image modality. For instance, TotalSegmentator has learnt a robust model for delineating 117 classes of objects in the human body, such as bones, a large subset of significant organs and veins [5]. Our application is unique because the PTV often includes a margin surrounding the visible tumour on the scan, which is different to other approaches which outline the boundaries of structures.

2.2 Transfer Learning

Transfer Learning uses knowledge that has been obtained from one task, and uses it as a starting point for learning a new task. It is therefore a useful solution to the problems identified in Section 2.1 because of the transferable knowledge features for similar domains and its proven success in generalizing features is trained properly.

The intuitive reason why transfer learning works is because in the early layers of deep learning, the model learns very low-level features. At this scale, the initial data-set or the cost function doesn't matter because a model working on the same problem but with different initialization will learn similar low-level features. This allows transferral because the (large/-sufficiently sized) input dataset is abstracted in the set low-level features which can instead be transferred. Then, the later layers are more specialized to a particular task [6]. It is similar to seeing the distribution in the training data change and transferring knowledge across domains [7].

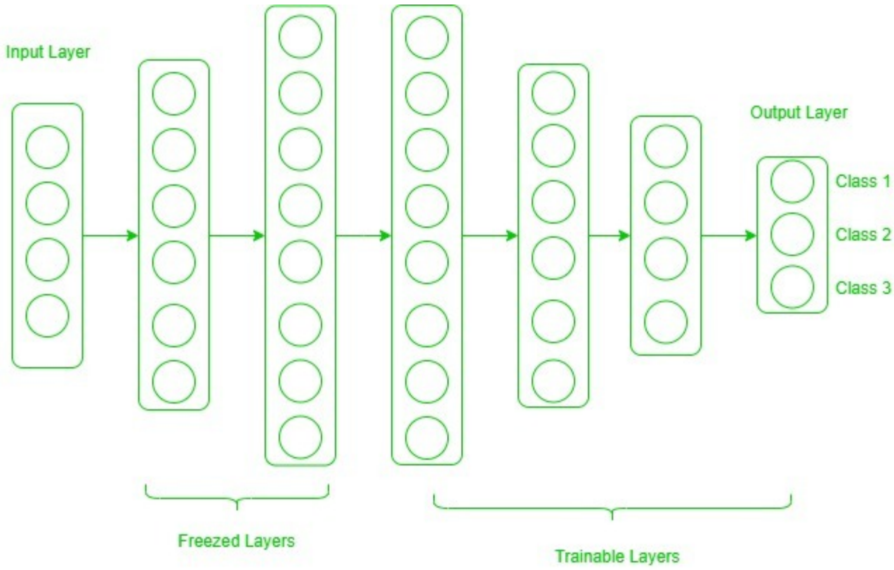

Figure 2.1: Early layers learn low-level features for similar domains, and during transfer of knowledge, these layers are frozen and the trainable layers are appended and weights are only updated for this layer [8]

Transfer Learning has the potential to: improve initial performance using only the transferred knowledge before nay further learning is done, improve the time it takes to fully learn the target task given the transferred knowledge, and improve the final performance all when compared to initial benchmarks without transfer [9]. It has also been found to work in medical contexts as well, where, for 332 abdominal liver CT scans, transfer learning generally improved weight initialization and resulted in faster convergence providing stronger and more robust representation [10]

Transfer Learning has been seen to prevent overfitting in domains where data volume is low and where generality without overfitting is hard to come by. This is because the model has already learnt features that are likely to be useful in the second task [8].

However, generalization is not a guarantee, as overfitting is still possible if the model is fine-tuned too much on the second task, as it may 'learn task-specific features that do not generalize well to new data' [8]. In our case, our target dataset is small, but similar to the base network dataset. Here, we may overfit because fine-tune the pre-trained network with the target dataset may not generalize to the global population. If instead we attempt to transfer a task with different base network dataset, then using high-level features of the pre-trained model will not be useful [8].

Bibliography

[1] William Small, Monica A. Bacon, Amishi Bajaj, Linus T. Chuang, Brandon J. Fisher, Matthew M. Harkenrider, Anuja Jhingran, Henry C. Kitchener, Linda R. Mileshkin, Akila N. Viswanathan, and David K. Gaffney. Cervical cancer: A global health crisis. *Cancer, An international interdisciplinary journal of American Cancer Society*, 123(13), 2017. URL https://acsjournals.onlinelibrary.wiley.com/doi/10.1002/cncr.30667. pages 1

[2] Zhikai Liu, Xia Liu, Bin Xiao, Shaobin Wang, Zheng Miao, Yuliang Sun, and Fuquan Zhang. Segmentation of organs-at-risk in cervical cancer ct images with a convolutional neural network. *Physica Medica*, 69:184–191, 2020. ISSN 1120-1797. doi: https://doi.org/10.1016/j.ejmp.2019.12.008. URL https://www.sciencedirect.com/science/article/pii/S1120179719305290. pages 1

[3] Lt Gen SR Mehta, Maj V Suhag, M Semwal, and Maj N Sharma. Radiotherapy : Basic concepts and recent advances. 2010. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4920949/pdf/main.pdf. pages 1

[4] Institute of Cancer Research and The Royal Marsden Hospital. Amlart data. pages 1

[5] wasserth, lassoan, fedorov, cnicolasgr, and Arputikos. URL https://github.com/wasserth/TotalSegmentator. pages 2

[6] Christopher M. Bishop and Hugh Bishop. *Deep Learning, Foundations and Concepts*. Springer, 2023. pages 3

[7] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. 2009. URL https://ieeexplore.ieee.org/abstract/document/5288526. pages 3

[8] What is transfer learning? URL https://www.geeksforgeeks.org/ml-introduction-to-transfer-learning/. pages 3

[9] Lisa Torrey and Jude Shavlik. Transfer learning. Technical report, University of Wisconsi, 2009. URL https://ftp.cs.wisc.edu/machine-learning/shavlik-group/torrey.handbook09.pdf. pages 3

[10] Michal Heker and Hayit Greenspan. Joint liver lesion segmentation and classification via transfer learning. Technical report, 2020. URL https://arxiv.org/pdf/2004.12352.pdf. pages 3