Imperial College London

MEng Individual Project

Department of Computing

Imperial College of Science, Technology and Medicine

Transfer Learning for Deep Learning Radiotherapy Planning

Author:
Anton Zhitomirsky

Supervisor:
Prof Ben Glocker

Second Marker:
Dr Thomas Heinis

May 29, 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 2.1.2). 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].

1.2 Current Solutions

The problem of automatic delineation of tumours in a patient is not a new concept unfamiliar to the scientific community. From 2016 to 2020 the number of deep-learning in radiotherapy publications has grown from 1,001 to 3,653 [5]. Therefore, there exist many proposals and solutions to this problem for cancerous tumours across the body. This also includes research projects focused specifically on cervical cancer and radiotherapy auto-contour planning. By conducting a accrual of literature across PubMed with the search string

```
"radiotherapy" AND "contour" AND ("cervix" OR "cervical") AND "cancer" AND ("Deep Learning" OR "
    ↪ DeepLearning" OR "Machine Learning" OR "ML" OR "Artificial Intelligence" OR "AI" OR "Computer
    ↪ assisted")
```

6 key papers published between the span of 2020 to 2021 were selected as most relevant for this project [6, 5, 7, 2, 8, 9]. These papers proposed novel ideas on how to solve image segmentation problems. Some noteworthy networks included Convolutional Neural Networks (Section **??**), 2D and 3D U-Net adaptations (Section **??** and **??**), V-Nets and DeepMedic models, with the most common approach being the U-Net approach according to a clinical survey on the topic in 2021 [6]. The methods in these papers will be further discussed in Chapter 2.

1.3 Motivation

Healthcare Crisis

Healthcare centres are experiencing high demand with low availability for all types of cancer. In England, waiting times are getting worse each month, with over a third of cancer patients (60,000) waiting beyond the 62-day target, and 10,000 patients waiting over 104 days according to a radiotherapy manifesto in 2022 *Manifesto link (shall i put in bibliography?)*.

This wait is partially due to the time consuming nature of treatment plans, with manual segmentation performed by a professional oncologist taking 90-120 minutes [2, 7] while always requiring optimization of radiotherapy placement to avoid oragns-at-risk [6].

Furthermore, more resources can be partitioned into other departments by developing software to aid with cancer treatments. A study found that the scale-up of radiotherapy capacity in 2015-35 from current levels 'could lead to saving of 26.9 million life-years in low-income and middle-income countries over the lifetime of the patients who received treatment' [10].

Current Tools

Unfortunately, there is a difference between auto-contouring and auto-segmentation. For our case we cannot use models such as TotalSegmentator (Section 2.2.1). TotalSegmentator has learnt form many samples to delineate 117 classes of objects in the human body, such as bones, a large subset of significant organs and veins [11]. Our case involves delineating a PTV, which is a significant challenge because it involves meta-information which is not visible on a CT scan, and therefore is not a trivial task of tracing around the visible contours of a tumour, but more so about adding sufficient padding to account for the microscopic spread of the tumour not visible to the scanner.

Since wrong or inaccurate contours constitute the highest factors for failure of treatment [8] current tools that are not bespoke for the task are therefore poor candidates for practical use.

Research Gap

The current literature makes great use of modern concepts stemming from the discovery of CNN models for auto-segmentation for radiotherapy planning. However, all papers assume a good sample size which represents the overall population of people under the same conditions. This is not practical in a general case, for instance: hospitals may only have the data for their in-house patients, teams operate on a niche areas and therefore have a smaller sample size, affected area scans are not standardized and vary between organizations which may promote the 'garbage in, garbage out' philosophy, and finally varying degrees of catastrophe levels pertaining to the sensitivity of the organs at risk (Section 2.1.3) surrounding the tumour (radiotherapy in the pelvis vs radiotherapy in the brain).

Summary

We therefore propose a solution that can be applied to groups with specific niche rules and small sample sizes. This Transfer Learning approach (Section 2.4) aims to leverage other trained models with great performance and lift low-level features to use in our case. This will also fill the literature gap which hasn't considered solving the auto-contouring for radiotherapy planning volumes, as a PubMed search querying publications with mentions of transfer learning returned no relevant matches.

1.4 Outline of Report

The structure of this report will take the reader of a reasonable programming background through the project such that they might be able to reconstruct the outcome themselves. It is expected of the reader to understand the concept of Machine Learning and an intuition surrounding Computer Vision.

Firstly Chapter 2 discusses the clinical context (Section 2.1) from which the idea for this project originates. We further construct the background knowledge for the current Computer Vision methodologies which exist to solve segmentation based issues (Section **??**). This will lead into currently pre-trained models (Section 2.2) and their application in a Transfer Learning setting (Section 2.4).

Chapter 2

Background

2.1 Clinical Context

This section will provide a baseline understanding of the clinical context for developing a tool to help segment tumours in patients with cervical cancer. Considering the unfortunate high frequency of cancers developing in people across the globe [1], the idea for developing a model to help segment cancers has long been in the scope for many researchers.

2.1.1 Cervical Cancer

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 2.1.2). This treatment is tedious, as it is estimated that an oncologist needs 90-120 min to delineate target areas for radiotherapy [2].

This time consuming nature of studying each patient makes treating cancers a time consuming endeavour which is particularly threatening when time is not on the patients side. Death from Cervical Cancer involves significant pain and suffering for the patients who cannot receive urgent treatment [1]. This is particularly a problem in mid-low income countries with approximately 85% of cases of cervical cancer occurring, making them have 18 times the death rates of high-income countries [1].

2.1.2 Radiotherapy

Radiation therapy is a treatment option for cancer treatment. In 2012, approximately 50% of all cancer patients received radiation therapy, with an additional 40% involving curative treatment [12].

Physicians abuse the physical properties of radiation to damage the genetic material of cells and block their ability to cause further damage to a patient [12]. Importantly for our cause, radiation damages intercellular molecules leading to degradation and stopping a cells ability to divide, leading to interphase death. Alternatively, radiation may cause a “mitotic catastrophe” causing a cell to die of a proliferative death [13].

2.1.3 Contour Planning

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.

Process

This begins with the high contrasting image of the tumour area. In a clinical setting, an Oncologist will use 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].

The first area defined from the scan is the Gross Target Volume (GTV). This macroscopic delineated area is the visible tumour area on the scan and contains a high probability of containing the tumour. Secondly, the Clinical Target Volume (CTV) is derived to account for potential microscopic spread. This will be an area at least as big as the tumor area segmented in the GTV with an optional margin surrounding it containing a 'rind' of non-zero probability of tumour spread. Lastly, the Primary Tumour Volume (PTV) contains residual geometric uncertainties and safety margins surrounding the CTV ensuring the radiotherapy dose is actually delivered to the CTV [14, 15, 5, 16]. These target volumes also constantly consider critical normal tissue structures which need to be preserved during irradiation. These are referred to as organs-at-risk (ORs). In some specific circumstances, it is necessary to add a margin analogous to the PTV margin around an OR to ensure that the organ cannot receive a higher-than-safe dose; this gives a planning organ at risk volume [15].

A tangent on Accuracy

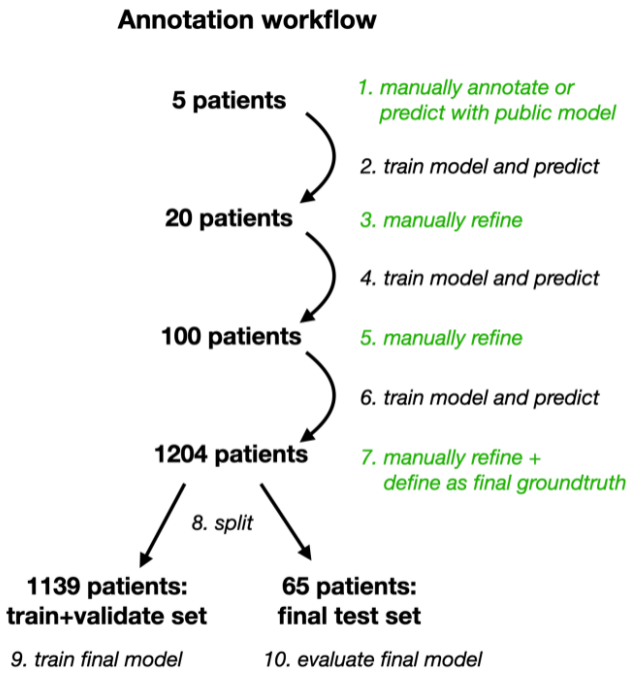
The CTV volumes may vary between clinicians as there is no internationally agreed guidelines. As a very time consuming process which high variability it therefore suffers from a lot of inter and intra-observer variability [5]. However, the data provided has been standardized as a gold standard, see Section **??**.

The PTV volume is an additional step which provides a margin of protection surrounding the CTV. One might think that a CTV step is the final step in clinical trials as it defines the most tight volume surrounding the cancerous site, however, geometric errors are impossible to eliminate. This includes short-term organ movement, voxel size and slice resolution, and possibility of relative movement of structure of reference and the tumour [17]

2.2 Existing Auto-Segmentation Methods

2.2.1 Total Segmentator

TotalSegmentator [18] is a model that was developed be a very robust segmentation tool. It has been trained on 1204 CT examinations and once more on 4004 whole-body CT examinations to investigate age dependent volume and attenuation changes. 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 [11]



It uses an iterative learning approach.

(1) After manual segmentation of the first 5 patients was completed, (2) a preliminary nnU-Net was trained, (3) and its predictions were manually refined, if necessary. (4) Retraining of the nnU-Net was performed after (5) reviewing and refining 5 patients, 20 patients, and (6) 100 patients.

In the end, all 1204 CT examinations had annotations that were manually reviewed and corrected whenever necessary. These final annotations served as the ground truth for training and testing. The model was trained on the dataset of 1082 patients, validated on the dataset of 57 patients and tested on the dataset of 65 patients. This final model was independent of the intermediate models trained during the annotation workflow, which reduced bias in the test set to a minimum. Using completely manual annotations in the test set would have introduced a distribution shift and thus greater bias [18].

TotalSegmentator is a model which encountered some issues with the dataset that may have impeded its performance. Firstly, some patients had ribs missing, which a clinician would typically count from top/bottom to identify them; however, on some scans these weren't visible, much like in our case, we are missing overies. These have been attributed to reasons for low performance for these structures, which warns us; our dataset contains many abnormalities according to the descriptions of each patient (Section **??**).

2.2.2 UniverSeg

2.2.3 SAM

TODO: maybe replace with medical SAM paper.

2.3 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.

2.3.1 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.

2.3.2 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 [11]. 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.

2.4 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 2.3 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 [19]. It is similar to seeing the distribution in the training data change and transferring knowledge across domains [20].

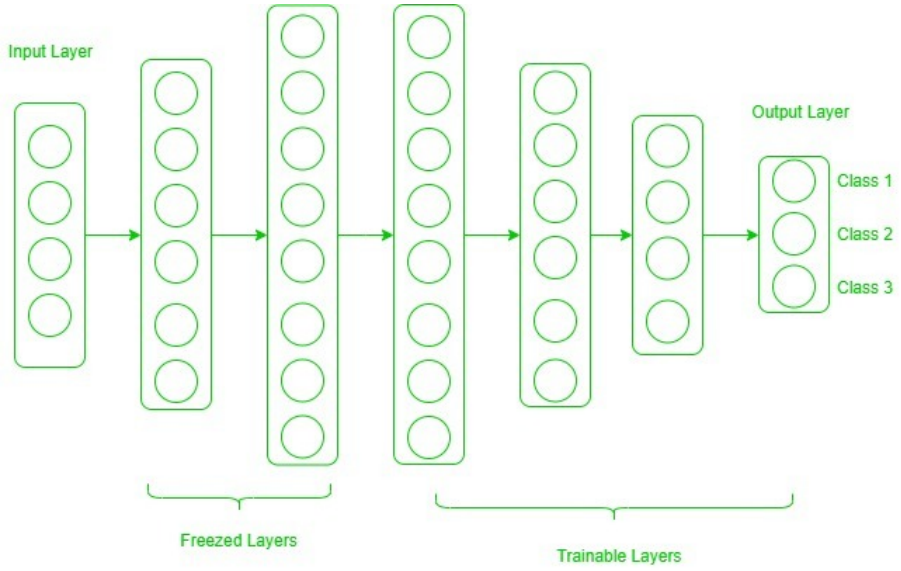


Figure 2.1: Early layers learn low-level features for similar domains, and during transfer of knowledge, these layers are frozen and the trainable layers are appended and weights are only updated for this layer [21]

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 [22]. 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 [23]

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 [21].

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' [21]. 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 [21].

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, 4

[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, 2, 4

[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, 5

[4] Institute of Cancer Research and The Royal Marsden Hospital. Amlart data. pages 1, 5

[5] Hui Lin, Haonan Xiao, Lei Dong, Kevin Boon-Keng Teo, Wei Zou, Jing Cai, and Taoran Li. Deep learning for automatic target volume segmentation in radiation therapy: a review. *Quant Imaging Med Surg*, 11(12):4847–4858, December 2021. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8611469/. pages 1, 2, 5

[6] Gihan Samarasinghe, Michael Jameson, Shalini Vinod, Matthew Field, Jason Dowling, Arcot Sowmya, and Lois Holloway. Deep learning for segmentation in radiation therapy planning: a review. *J Med Imaging Radiat Oncol*, 65(5):578–595, July 2021. URL https://pubmed.ncbi.nlm.nih.gov/34313006/. pages 2

[7] Hanna Sartor, David Minarik, Olof Enqvist, Johannes Ulén, Anders Wittrup, Maria Bjurberg, and Elin Trägårdh. Auto-segmentations by convolutional neural network in cervical and anorectal cancer with clinical structure sets as the ground truth. *Clin Transl Radiat Oncol*, 25:37–45, September 2020. URL https://pubmed.ncbi.nlm.nih.gov/33005756/. pages 2

[8] Dong Joo Rhee, Anuja Jhingran, Bastien Rigaud, Tucker Netherton, Carlos E Cardenas, Lifei Zhang, Sastry Vedam, Stephen Kry, Kristy K Brock, William Shaw, Frederika O'Reilly, Jeannette Parkes, Hester Burger, Nazia Fakie, Chris Trauernicht, Hannah Simonds, and Laurence E Court. Automatic contouring system for cervical cancer using convolutional neural networks. *Med Phys*, 47(11):5648–5658, October 2020. URL https://pubmed.ncbi.nlm.nih.gov/32964477/. pages 2

[9] Zhikai Liu, Xia Liu, Hui Guan, Hongan Zhen, Yuliang Sun, Qi Chen, Yu Chen, Shaobin Wang, and Jie Qiu. Development and validation of a deep learning algorithm for auto-delineation of clinical target volume and organs at risk in cervical cancer radiotherapy. *Radiotherapy and Oncology*, 153:172–179, 2020. ISSN 0167-8140. doi: https://doi.org/10.1016/j.radonc.2020.09.060. pages 2

[10] Prof Rifat Atun, Prof David A Jaffray, Michael B Barton, Freddie Bray, Prof Michael Baumann, Bhadrasain Vikram, Timothy P Hanna, Felicia M Knaul, Yolande Lievens, Tracey Y M Lui, Prof Michael Milosevic, Prof Brian O'Sullivan, Danielle L Rodin, Eduardo Rosenblatt, Prof Jacob Van Dyk, and Mei Ling Yap. Expanding global access to radiotherapy. *The Lancet Oncology Commission*, 16(10), 2015. URL https://www.thelancet.com/journals/lanonc/article/PIIS1470-2045(15)00222-3/fulltext. pages 2

[11] wasserth, lassoan, fedorov, cnicolasgr, and Arputikos. URL https://github.com/wasserth/TotalSegmentator. pages 2, 5, 7

[12] Rajamanickam Baskar, Kuo Ann Lee, Richard Yeo, and Kheng-Wei Yeoh1. Cancer and radiation therapy: Current advances and future directions. 2012. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3298009/. pages 4

[13] Yunfei Jiao, Fangyu Cao, and Hu Liu. Radiation-induced cell death and its mechanisms. *Halth Phys.*, 2022. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9512240/pdf/hpj-123-376.pdf. pages 4

[14] C. F. Njeh. Tumor delineation: The weakest link in the search for accuracy in radiotherapy, 2008. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2772050/. pages 5

[15] Neil G Burnet, Simon J Thomas, Kate E Burton, and Sarah J Jefferies. Defining the tumour and target volumes for radiotherapy, 2004. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1434601/pdf/ci040153.pdf. pages 5

[16] David Bernstein, Alexandra Taylor, Simeon Nill, and Uwe Oelfke. New target volume delineation and ptv strategies to further personalise radiotherapy, 2021. pages 5

[17] Marcel van Herk. Errors and margins in radiotherapy. *Seminars in Radiation Oncology*, 14(1):52–64, 2004. ISSN 1053-4296. doi: https://doi.org/10.1053/j.semradonc.2003.10.003. URL https://www.sciencedirect.com/science/article/pii/S1053429603000845. High-Precision Radiation Therapy of Moving Targets. pages 5

[18] Jakob Wasserthal, Hanns-Christian Breit, Manfred T. Meyer, Maurice Pradella, Daniel Hinck, Alexander W. Sauter, Tobias Heye, Daniel T. Boll, Joshy Cyriac, Shan Yang, Michael Bach, and Martin Segeroth. Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images. 2023. URL https://arxiv.org/pdf/2208.05868.pdf. pages 5, 6

[19] Christopher M. Bishop and Hugh Bishop. *Deep Learning, Foundations and Concepts*. Springer, 2023. pages 7

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

[21] What is transfer learning? URL https://www.geeksforgeeks.org/ml-introduction-to-transfer-learning/. pages 7, 8

[22] 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 8

[23] 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 8