

CONE-BEAM COMPUTED TOMOGRAPHY (CBCT) SEGMENTATION BY
ADVERSARIAL LEARNING DOMAIN ADAPTATION

A Thesis

by

XIAOQIAN JIA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Chair, Xiaoning Qian Co-
Co-Chair of Committee, Jim Ji
Committee Members, Ulisses Braga Neto
Tie Liu
Xia Hu
Head of Department, Miroslav Begovic

December 2019

Major Subject: Electrical Engineering

Copyright 2019 Xiaoqian Jia

ABSTRACT

Cone-beam computed tomography (CBCT) is increasingly used in radiotherapy for patient alignment and adaptive therapy where organ segmentation and target delineation are often required. However, due to the poor image quality, low soft tissue contrast, as well as the difficulty in acquiring segmentation labels on CBCT images, developing effective segmentation methods on CBCT has been a challenge. In this thesis, we propose a deep model for segmenting organs in CBCT images without requiring labelled training CBCT images.

By taking advantage of the available segmented computed tomography (CT) images, our adversarial learning domain adaptation method aims to synthesize CBCT images from CT images. Then the segmentation labels of the CT images can help train a deep segmentation network for CBCT images, using both CTs with labels and CBCTs without labels. Our adversarial learning domain adaptation is integrated with the CBCT segmentation network training with the designed loss functions. The synthesized CBCT images by pixel-level domain adaptation best capture the critical image features that help achieve accurate CBCT segmentation. Our experiments on the bladder images from Radiation Oncology clinics at the University of Texas Southwestern Medical School (UTSW) have shown that our CBCT segmentation with adversarial learning domain adaptation significantly improves segmentation accuracy compared to the existing methods without doing domain adaptation from CT to CBCT.

DEDICATION

To my parents, my grandparents, and my best friend and family, Dongzuo.

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Xiaoning Qian for the continuous support of my master study and research, for his patience, motivation, enthusiasm, and immense knowledge. He helps me in all the time of research and writing of this thesis. Only with his advice and guidance that I can go far in my field of study. Besides my advisor, I would like to thank my co-chair Prof. Jim Ji and the rest of my thesis committee members, insightful comments, and hard questions. I also want to thank Prof. Steve Jiang from UT southwestern Medical School, who supported me and organized this project.

My sincere thanks also go to my lab-mate, Sicheng Wang. I will always remember, every midnight we work in the lab, the hardship and challenges we meet in the project, and every moment we spend in Dallas. Moreover, I shall thank for my love, Dongzuo Tian, without his spiritual support, I cannot complete such a excellent job. Last but not the least, I would like to thank my family: my parents Hongyun Fu and Jianzeng Jia, for giving birth to me at the first place and supporting me spiritually throughout my life.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a thesis committee consisting of Professor Xiaoning Qian, Professor Jim Ji, Professor Ulisses Braga-Neto and Professor Tie Liu of the Department of Electrical and Computer Engineering Department and Professor Xia Hu of the Department of Computer Science and Engineering.

The CBCT data was provided by Professor Steve Jiang from UT Southwestern Medical School.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

Graduate study was supported by Graduate Merit Scholarship from Texas A&M University and a dissertation research fellowship from Medical Artificial Intelligence and Automation Laboratory, Department of Radiation Oncology, University of Texas Southwestern, USA.

NOMENCLATURE

CBCT	Cone beam computed tomography
CT	Computed tomography
ART	Adaptive Radiotherapy
DA	Domain Adaptation
GAN	Generative Adversarial Networks
CycleGAN	Cycle Generative Adversarial Networks
Seg	Segmentation
CNN	Convolutional Neural Network
DSC	Dice Coefficient Score

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGMENTS	iv
CONTRIBUTORS AND FUNDING SOURCES	v
NOMENCLATURE	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	viii
LIST OF TABLES.....	ix
1. INTRODUCTION.....	1
2. LITERATURE REVIEW	5
2.1 CBCT Application in Adaptive Radiotherapy (ART)	5
2.2 Deep Segmentation	6
2.3 Generative Adversarial Network	7
2.4 Domain Adaptation	9
2.5 Deep Learning Applications in Medical Imaging	10
3. MODEL FORMULATION	11
3.1 Adversarial Learning Domain Adaptation	12
3.2 Deep Segmentation via Adversarial Learning	16
3.3 Learning Diagram	18
4. EXPERIMENTS AND RESULTS	20
4.1 Datasets	20
4.2 Implementation Details	20
4.3 Experimental Results & Discussion.....	21
5. CONCLUSION AND FUTURE WORK	28
REFERENCES	30

LIST OF FIGURES

FIGURE	Page
1.1 ART using CBCT	2
1.2 Examples of wrong CBCT contours from DIR.	3
2.1 Residual learning: a building block	7
2.2 Architecture of GAN.....	8
3.1 Workflow of our proposed method.	12
3.2 The CycleGAN architecture.	13
3.3 U-net	15
3.4 CycleGAN discriminator architecture.....	15
3.5 Segmentation module	17
3.6 Feature discriminator architecture.....	18
4.1 Visualization of example CT images	24
4.2 Visualization of example CBCT images	25
4.3 Domain adaptation between CT and CBCT.....	26
4.4 Visual comparison of segmentation results by different methods.....	27

LIST OF TABLES

TABLE	Page
4.1 Performance comparison for CBCT bladder segmentation.	22
4.2 Sensitivity analysis from λ_{seg} 1 to 10	23

1. INTRODUCTION

Radiotherapy is one of the most effective ways to treat cancer. Cancer radiotherapy could take several weeks. During the process, patient's anatomy may change significantly and the initially optimized treatment plan may become sub-optimal, leading to degraded treatment outcome. One way to address this problem is adaptive radiation therapy (ART) [1], where the treatment plan is re-optimized using the updated patient anatomy right before the treatment on a particular treatment day. Cone-beam computed tomography (CBCT), the most widely available 3D imaging modality on modern linacs, is commonly used for ART re-planning, whose efficacy depends on accurate segmentation of the organs and treatment target(s) in CBCT images.

The typical workflow of ART is shown in Fig. 1.1. Patients have CT scans when initializing the treatment. With the planning CT scans, physicians draw contours and calculate the dose prediction based on them to make a treatment plan. In the traditional treatment, the plan would keep the same without considering the tumor shrinking or anatomy changing. The aim of ART is to adapt the radiotherapy by using CBCT to observe the patients' organ and tumor changes during the treatment. These changes should be considered to adjust the radiotherapy location and dose. For example, if the tumor becomes larger or smaller, which means the current plan is not suitable, it is necessary to let the patient re-plan again according to the updated record. Obtaining precise CBCT contours/segmentation is the key to the success of ART.

The main challenge when analyzing CBCT for ART is, unlike CT images, whose image quality is much better and the segmentation labels are readily available from the routine treatments, accurate segmentation in CBCT images is far more challenging. But, CBCT is widely used for patient positioning in radiation therapy. It has the most up-to-date patient anatomy as its frequently taken. But due to inaccurate Hounsfield Unit (HU) values, more artifacts and scatters on CBCT compared with CT, its applications are limited. Thus, labels are not part of routine clinical work, and therefore are not available in a supervised learning setting.

The popular way to analyze CBCT images for ART is deformable image registration (DIR) [2].

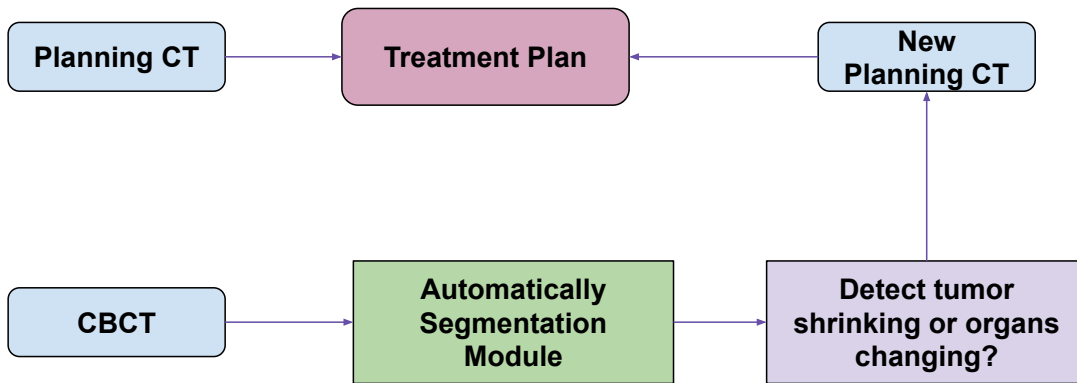


Figure 1.1: ART using CBCT

It is to register the planning CT (pCT) with daily CBCT, also register the pCT contours of multi-organs into CBCT. However, since the CBCT images can have some scatter, noise, artifacts [3] and inaccurate HU values, the DIR sometimes is problematic. The DIR can be wrong because the CBCT, and CT exists a significant difference between their anatomy. Due to this, the deformed CT (dCT) images sometimes will not represent the anatomy in CBCT well. As shown in Fig. 1.2, the orange contours from DIR are also inaccurate, and cannot be used in ART.

In the Computer Vision and Medical Imaging research, there are many existing pipelines in image domain adaptation and segmentation. Recently, the generative modules, such as generative adversarial networks (GANs) [4] and variational autoencoders (VAEs) [5], aim to generate synthesized images similar as the ones from the input domain of interest. This approach has been applied to medical image analysis in [6], for synthesizing unpaired head and neck MR and CT images. For CT and CBCT datasets, the authors in [7] have explored image translation between Head&Neck CT and CBCT dataset using a CycleGAN [8] module. Fully-convolutional Networks (FCNs) [9] have been applied for medical image segmentation. For example, [10] implemented the 3D FCN in male pelvic CT images, to segment prostate and surrounding organs at risk (OARs). The authors in [11] and [12] combined image domain adaptation and FCN segmentation to segment cardiac CT images and obtain the improved results.

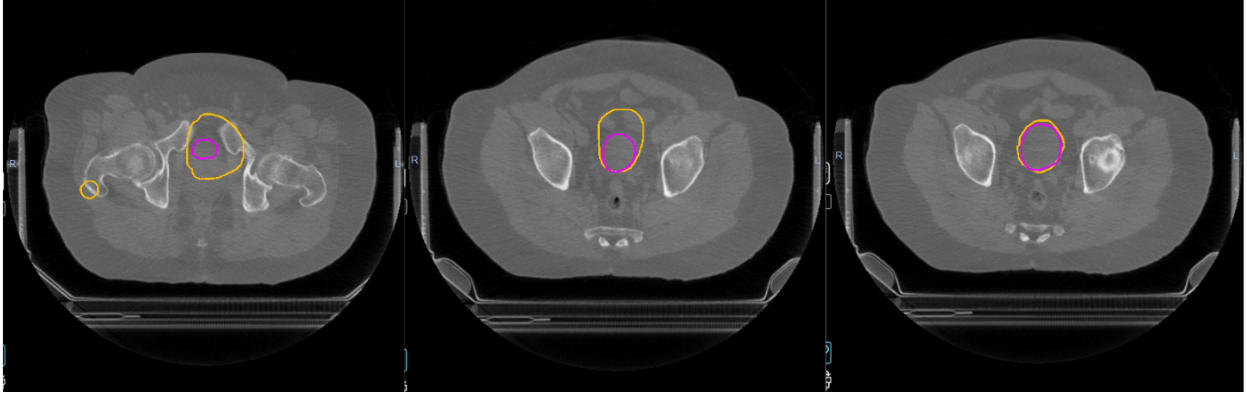


Figure 1.2: Examples of wrong CBCT contours from DIR: In the images, the orange contour is from Deformable registration algorithm by a open-source software Velocity AI, and the pink one is from a professional physician.

Due to the lower quality, wrong HU value, lack of annotated organ contours, we cannot do CBCT segmentation in the naive supervised setting. In this thesis, we propose to solve this problem by developing an adversarial learning domain adaptation, aiming to train a segmentation without using CBCT’s contours. Specifically, we address the following two main questions: 1. How to apply recent deep learning methods in CBCT image segmentation without training segmentation labels? 2. Since in clinical settings, patients in ART have their preoperative CT images with labeled organ contours. Can we utilize these labels to help CBCT segmentation task?

First, we plan to use GAN to convert the preoperative CT images to CBCT-like images, named sCBCT (synthetic CBCT), and use sCBCT and its corresponding annotated CT contours to train segmentation in the supervised learning setting. At the same time, we will perform adversarial learning for pixel-level domain adaptation between sCBCT and CBCT images. Eventually, the adapted model translates CT contours for CBCT images. The algorithm is designed using two deep neural networks, i.e., pre-trained CycleGAN transformation and FCN image segmentation. We use the output from the segmentation model to train a pixel-level discriminator, which will leverage pre-trained sCBCT segmentation labels into the CBCT domain to guide CBCT image segmentation.

In Chapter 2, we will review the literature on CBCT application in ART, image segmentation

and domain adaptation. Chapter 3 introduces our proposed method, including the network architecture and loss function definitions. The implement details, evaluation results and discussion are described in Chapter 4. Finally, Chapter 5 concludes the thesis and describes potential future research directions.

2. LITERATURE REVIEW

Our purpose is to generate CBCT contours via advanced machine learning methods to guide CBCT segmentation. To get the reasonable CBCT contours, transfer learning or domain adaptation between CT and CBCT images can be adopted to help train robust image transform and segmentation modules jointly for reliable segmentation of CBCT images, on which medical physicians can carry out better dose prediction and adaptive radiation therapy planning respectively.

In this chapter, we will first review recent works in CBCT application for ART, deep segmentation, generative adversarial network, and domain adaptation, which contribute to different components of our proposed CBCT segmentation model.

2.1 CBCT Application in Adaptive Radiotherapy (ART)

In the radiation therapy, to confirm the correctness of the treatment delivery, it needs a dose guidance procedure. Essentially, with visualisation of the dose distribution in three-dimensional images (e.g. CBCT) created just before the dose delivery, the dose guidance procedure can estimate the collection of radiation (by a detector) that is actually delivered to the patient.

Currently, re-planning doses based on images collected during the treatment is regarded as the state-of-the-art for dose-guided procedures [13, 14]. Non-rigid/deformable registration and dose mapping algorithms for dose accumulation should cause some anatomy changes in different fractions. Based on anatomy/geometry changes and the dose delivered to tumor, it is often required to check the agreement between the treatment plan and its implementation. If there is any disagreement, the initial treatment plan should be re-planned to the clinically confident one to meet the radiation therapeutic objectives. This is called adaptive radiation therapy (ART). Thus, to utilize CBCT in ART to track possible changes of the organ contours between CT and CBCT, different registration methods have been proposed. For example, Li *et al.* [15] adopted rigid-registration in a hybrid approach called “AIGRT” (adaptive image guided radiation therapy). The authors in [16, 17, 18, 19] proposed deformable methods to map HU values between CBCT and CT images.

Boggula *et al.* [20] developed an algorithm using multi-level threshold values for HU conversion from CT to CBCT images.

2.2 Deep Segmentation

A lot of famous convolutional neural networks (CNNs) have been developed. LeNet [21] has demonstrated successful digit image classification. Then, AlexNet[22] has been released, which is a deeper and wider version of the LeNet, to learn more complex objects in the image. Szegedy *et al.* proposed GoogleNet [23], a combination of convolutional filters and pooling layers in its main architecture. Then, Inception v2-v4 were developed [24, 25] with some modifications to GoogleNet. He *et al.* [26] proposed ResNet, using a new way to solve the vanishing gradient problem by skipping one or more layers, as shown in Fig. 2.1. The gradients training ResNet can easily flow to the shortcut connection without any obstacle during back propagation. Although these networks are successful in image classification on famous benchmark datasets (MNIST digit, ImageNet, etc.), they have limited performance for image segmentation. Because of the fully connected layers in each architecture, these CNNs can only be applied to segment small images or small patches of given images.

Long *et al.* [9] proposed a fully-convolutional network (FCN) framework, which replaces all the fully connected layer with convolutional layers in the popular CNNs mentioned before and can transfer them for effective image segmentation. The output of these networks are the spatial heatmaps and the deconvolution layers for upsampling enabled dense inference and learning the per-pixel labels for segmentation. Although FCN is successful, it does not take into account useful global context information. SegNet [27] is a successful example to introduce an encoder-decoder structure in FCN to further improve segmentation performance. The decoder in SegNet is constructed by a set of upsampling deconvolution layers followed by a softmax output layer to predict the pixel-wise labels.

In medical imaging, Olaf Ronneberger *et al.* [28] released Unet, which is also based on the encoder-decoder architecture. Unet contains two parts, the encoder (or contraction path) is a traditional stack of convolution layers and maxpooling layers, to capture the context in images. The

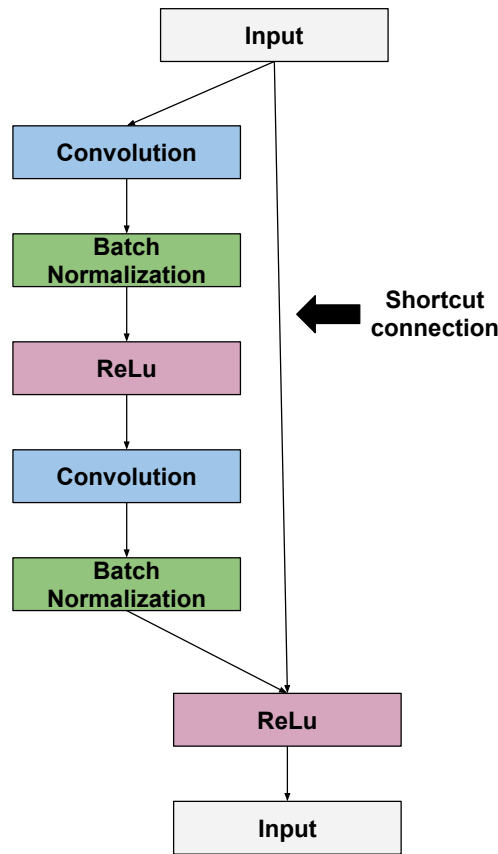


Figure 2.1: Residual learning: a building block

decoder is the symmetric expanding path using transposed convolution (deconvolution) layers, to enable more accurate localization. Unet has played a significant role in medical image segmentation. For example, the authors in [29] applied a dense Unet to liver and tumor segmentation in CT images. In [30], a modified Unet by adding more intermediate layers to skip connections allows more new upsampling paths from different depths, enables various receptive fields to better capture global and local information to achieve accurate segmentation.

2.3 Generative Adversarial Network

In the past few years, generative adversarial network (GAN) has gained more attention in computer vision because of their significant performance in image generation, image to image transla-

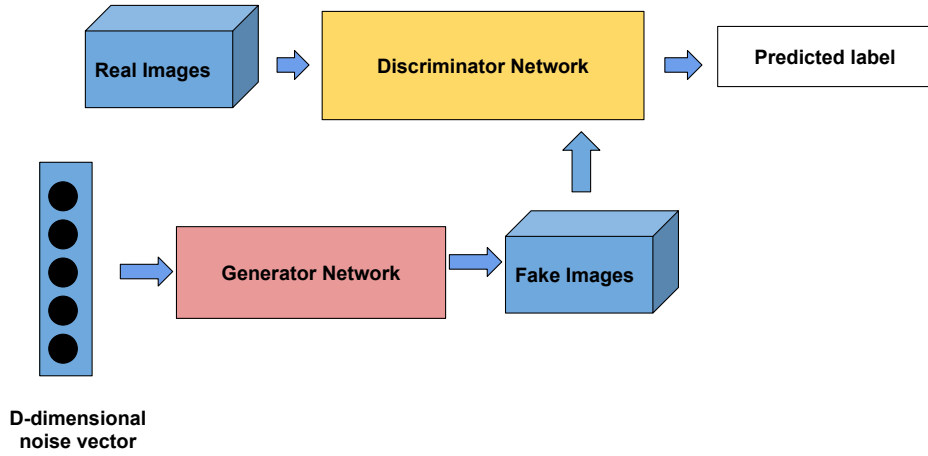


Figure 2.2: Architecture of GAN

tion, etc. I. Goodfellow *et al.* [4] proposed the original GAN in 2014.

The structure of the original GAN, shown in Fig. 2.2, has two components: one is a discriminator (D) to distinguish between real images and generated images, while the other one is a generator (G) to generate images to fool the discriminator. Give a distribution modeling real images $x \sim \mathcal{P}_{data}$ and input noise variables $p_z(Z)$, G models a probability distribution \mathcal{P}_g over data x . The loss function to help train D and G is:

$$L_{GAN} = \min_G \max_D E_{x \sim \mathcal{P}_{data}(x)} [\log D(x)] + E_{z \sim \mathcal{P}_z(z)} [\log(1 - D(G(z)))]. \quad (2.1)$$

More extensions of GAN have been proposed, DCGAN [31] has added a deconvolutional layer to the generator to enable the generation of higher resolution images. Progressive GAN (PROGAN) [32] updates the network architecture in progressive steps with training starting with 4×4 small image and then growing the training image size. Recently, BigGAN [33] has also been designed to improve performance by increasing model complexity and batch size. CycleGAN [8]

and pixel-to-pixel GAN [34] have been proposed for image translation between two domains. The latter requires that images from two domains should be paired while CycleGAN doesn't require the paired data. CycleGAN also uses the cycle-consistent loss as one part of its training objective function, to make sure about the image style transformation. It has been widely used in medical image translations between two modalities.

2.4 Domain Adaptation

Domain adaptation (DA) in deep learning has been viewed as a new strategy to solve the learning challenge due to the lack of labeled data. In the real-world applications, it can be extremely expensive and time-consuming to get enough labelled data. But, other domains and tasks may have enough number of data. So, using data from other domain(s) may help the current task, which doesn't have enough training data. However, there is always some domain shift or distribution gap between two domains. If using a learning model trained for one domain to another domain directly, it may degrade the performance. The purpose of domain adaptation is to find an "adaptation" method to let the model from one domain to work well in another domain.

The most straightforward way to do domain adaptation is fine-tuning the model. There have been some methods trying to minimize the statistical distribution shift. Long *et al.* [35] proposed Deep Adaptation Network (DAN) to use the maximum mean discrepancy (MMD) in the training objective function to reduce the domain discrepancy. Zhuang *et al.* [36] proposed to minimize the Kullback-Leibler (KL) divergence for DA. With the development in GAN, adversarial models were shown to achieve significant improvement in DA. Tzeng *et al.* [37] proposed adversarial discriminative domain adaptation (ADDA), using an additional domain classification between source and target domains. Cao *et al.* [38] released a selective adversarial network (SAN), to obtain partial DA from large domain to small domain. Based on Wasserstein GAN [39], Shen *et al.* [40] used a discriminator to estimate the Wasserstein distance between source and target domain, and optimize the network to minimize the distance. Hoffman *et al.* [41] proposed a framework that combines CycleGAN and ADDA, to derive both the pixel-level and feature-level DA.

2.5 Deep Learning Applications in Medical Imaging

Recently, deep learning modules have been widely applied in medical image analysis, especially in image segmentation. Accurate segmentation of OAR (organs at risk) in medical images allows more precise quantitative analysis of organ's size and shape, as in cancer treatment. Besides Unet mentioned in the previous section, Milletari *et al.* [42] proposed a Vnet that is the 3D extension of Unet. Balagopal *et al.* [10] applied a 3D CNN network for male pelvic CT datasets. There have been also a lot of works based on GAN for medical image reconstruction. For example, Ran *et al.* [43] used wGAN for 3D MRI denoising. And Oksuz *et al.* [44] did the artifact correction by GAN. Recently, for domain adaptation in medical imaging, Chen *et al.* [45] released an adversarial domain adaptation model between two Xray datasets. They *et al.* [12] also proposed a similar idea for MRI and CT images. Dou *et al.* [46] developed a DA method based on a plug-and-play framework between cardiac MRI and CT images.

3. MODEL FORMULATION ¹

As we discussed, it is hard to achieve accurate organ segmentation for CBCT images due to their lower image quality. Compared with CT, CBCT has inaccurate Hounsfield Unit (HU) values, more artifacts and scatters. The traditional deformable image registration (DIR) methods to infer CBCT segmentation by deforming the CT anatomy have limited successes because of significant appearance differences across two modalities. In addition, it is infeasible to acquire manual segmentation labels for CBCT images. Hence, for adaptive treatment planning using CBCT, neither the traditional image registration nor supervised image segmentation using deep networks is effective.

To address the challenges in CBCT segmentation, we propose to first use CycleGAN to bridge the domain gap between CT and CBCT, and to apply the adversarial learning into segmentation training. The proposed framework consists of two parts, image domain adaptation and segmentation. We adopt model training based on the CyCADA architecture [41]. CyCADA enables task-driven adversarial learning by combining domain adaptation using CycleGAN [8] with image classification or segmentation as the ultimate goal. Specifically, CycleGAN is constituted by two generative adversarial networks (GANs) [4] to transfer images between domains through a consistency loss function, requiring no paired data when training. The integration of adversarial domain adaptation and task-driven adversarial learning in CyCADA can capture both pixel-level and feature-level domain invariant representations and therefore better helps the ultimate task.

In this chapter, we introduce the proposed adversarial domain adaptation guided image segmentation network. As illustrated in Fig. 3.1, our network consists of two critical modules: adversarial domain adaptation and deep segmentation. These two modules are intertwined with the segmentation module designed to provide necessary anatomical details as the feedback to help better guide the synthesis of CBCT images from training CT images in adversarial domain adaptation. We will describe these components in detail in the following two subsections.

¹Part of this section is reprinted from our accepted manuscript [47] ©2019 MICCAI

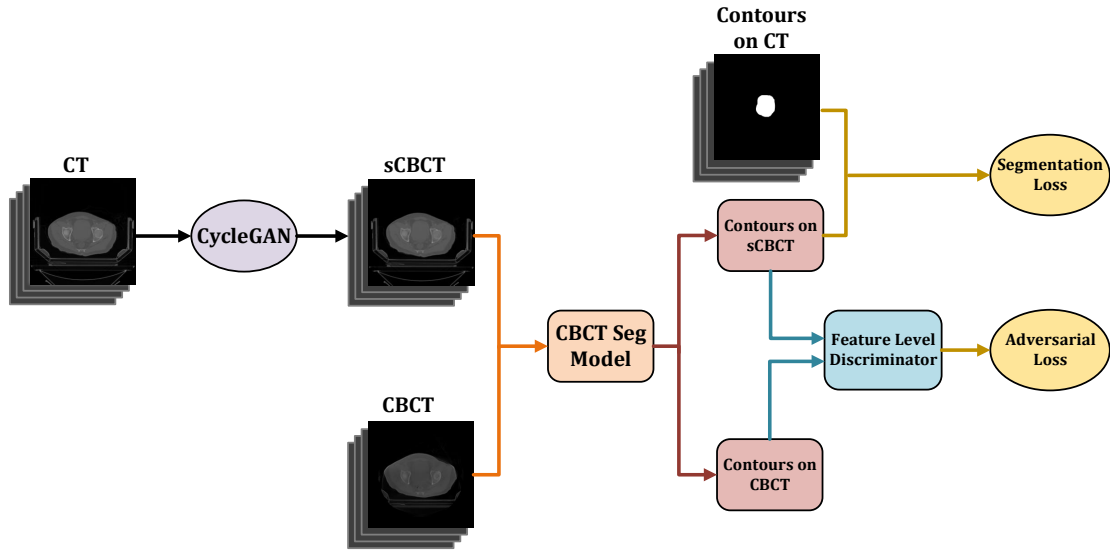


Figure 3.1: Workflow of our proposed method.

3.1 Adversarial Learning Domain Adaptation ²

We first introduce the adversarial domain adaptation module based on CycleGAN. The key idea here is to develop a generative model for effective synthesis of CBCT images with inherited CT image segmentation labels so that labeled CT images can be used for CBCT segmentation training.

In our implementation of CycleGAN for adversarial domain adaptation (in Fig. 3.2), CycleGAN has two generators G_{CT} and G_{CBCT} , which synthesize CT and CBCT images respectively. The synthesized images will be judged by two corresponding discriminators D_{CT} and D_{CBCT} . During the training, the synthesized images will be compared to the corresponding CT images in the source domain and CBCT images in the target domain. The generators also derive the “CycleCT” and “CycleCBCT” images from the synthesized images in a cyclic fashion. These images will also be compared to the original CT and CBCT images to achieve “cyclic consistent” domain adaptation.

In this framework, by training CycleGAN, two generators aim to synthesize the CT/CBCT images based on training CBCT/CT images so that two discriminators can not distinguish between

²Part of this section is reprinted from our accepted manuscript [47] ©2019 MICCAI

synthesized images and original training images. Moreover, CycleGAN enforces the cyclic consistency to make sure that the images synthesized by two generators in a cyclic fashion have the consistent quality in the sense that generators can fool discriminators into believing that the generated “fake” images are “real”.

For training the discriminator and generator for synthesized CT (sCT) images, the loss function is:

$$L_{GAN}(G_{CT}, D_{CT}) = E_{x \sim \mathcal{P}_{CBCT}(x)} [D_{CT}(G_{CT}(x))^2] + E_{y \sim \mathcal{P}_{CT}(y)} [(1 - D_{CT}(y))^2], \quad (3.1)$$

where \mathcal{P}_{CT} and \mathcal{P}_{CBCT} are the sets of unpaired input training CT and CBCT images. Again, the generator G_{CT} aims to generate sCT images that are similar to the input CT images; and D_{CT} aims to distinguish between sCT images and original input CT images \mathcal{P}_{CT} . Similarly, for the synthesized CBCT (sCBCT) branch in Fig. 3.2 (top), we have the corresponding MSE-based adversarial learning loss:

$$L_{GAN}(G_{CBCT}, D_{CBCT}) = E_{y \sim \mathcal{P}_{CT}(y)} [D_{CBCT}(G_{CBCT}(y))^2] + E_{x \sim \mathcal{P}_{CBCT}(x)} [1 - D_{CBCT}(x)]^2. \quad (3.2)$$

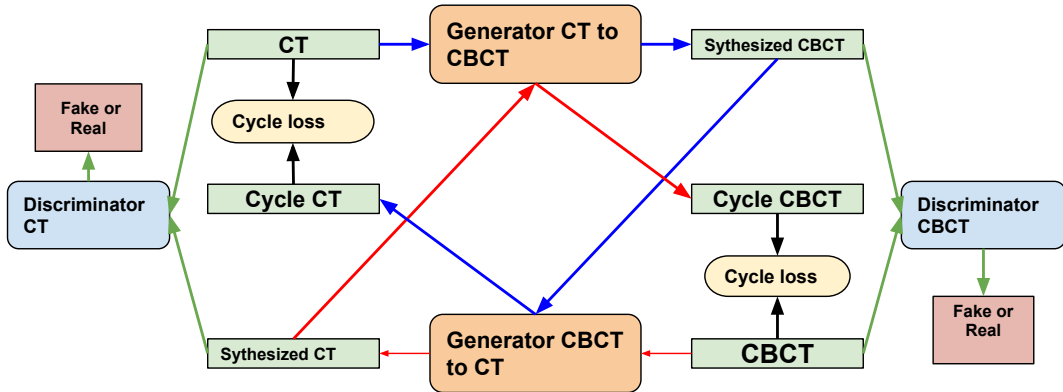


Figure 3.2: The CycleGAN architecture is used to generate sCBCT images from CT images.

To guarantee the high-quality domain adaptation, an additional cycle consistency loss [8] is imposed to have the reconstructed image in a cyclic fashion become identical to the original input training images. This is done by imposing an \mathcal{L}_1 difference on the reconstruction error, given as:

$$\begin{aligned} L_{cycle}(G_{CT}, G_{CBCT}) = & E_{y \sim \mathcal{P}_{CT}(y)} [\| (G_{CT}(G_{CBCT}(y)) - y) \|_1] \\ & + E_{x \sim \mathcal{P}_{CBCT}(x)} [\| (G_{CBCT}(G_{CT}(x)) - x) \|_1]. \end{aligned} \quad (3.3)$$

Furthermore, an identity loss is added:

$$\begin{aligned} L_{identity}(G_{CT}, G_{CBCT}) = & E_{y \sim \mathcal{P}_{CT}(y)} [\| (G_{CT}(y) - y) \|_1] \\ & + E_{x \sim \mathcal{P}_{CBCT}(x)} [\| (G_{CBCT}(x) - x) \|_1], \end{aligned} \quad (3.4)$$

leading to the final total loss function:

$$\begin{aligned} L_{cyclegan} = & L_{GAN}(G_{CT}, D_{CT}) + L_{GAN}(G_{CBCT}, D_{CBCT}) \\ & + \lambda_{cycle} L_{cycle}(G_{CT}, G_{CBCT}) + \lambda_{id} L_{identity}(G_{CT}, G_{CBCT}), \end{aligned} \quad (3.5)$$

where λ_{cycle} and λ_{id} controls the relative importance of the L_{cycle} and $L_{identity}$. Notice the symmetry of the designed loss functions to have the cyclic consistency guarantee for better quality adversarial domain adaptation. For CBCT segmentation, we focus on the sCBCT branch on the top of Fig. 3.2. When the training reaches an optimum with respect to the total loss function, we hope that from the source CT images, we can generate sCBCT images whose representations are similar to those of the target CBCT images so that they can be used to train an effective CBCT segmentation network.

In our implementation, as shown in Fig. 3.3, we choose the generator network architecture to be the U-net [28] for end-to-end pixel-level image transformation, as similarly adopted in the original CycleGAN [8], where the inputs and outputs of the U-net are $512 \times 512 \times 1$ images. For discriminators, as shown in Fig. 3.4, a 142×142 PatchGAN [34] is applied to output the $32 \times 32 \times 1$

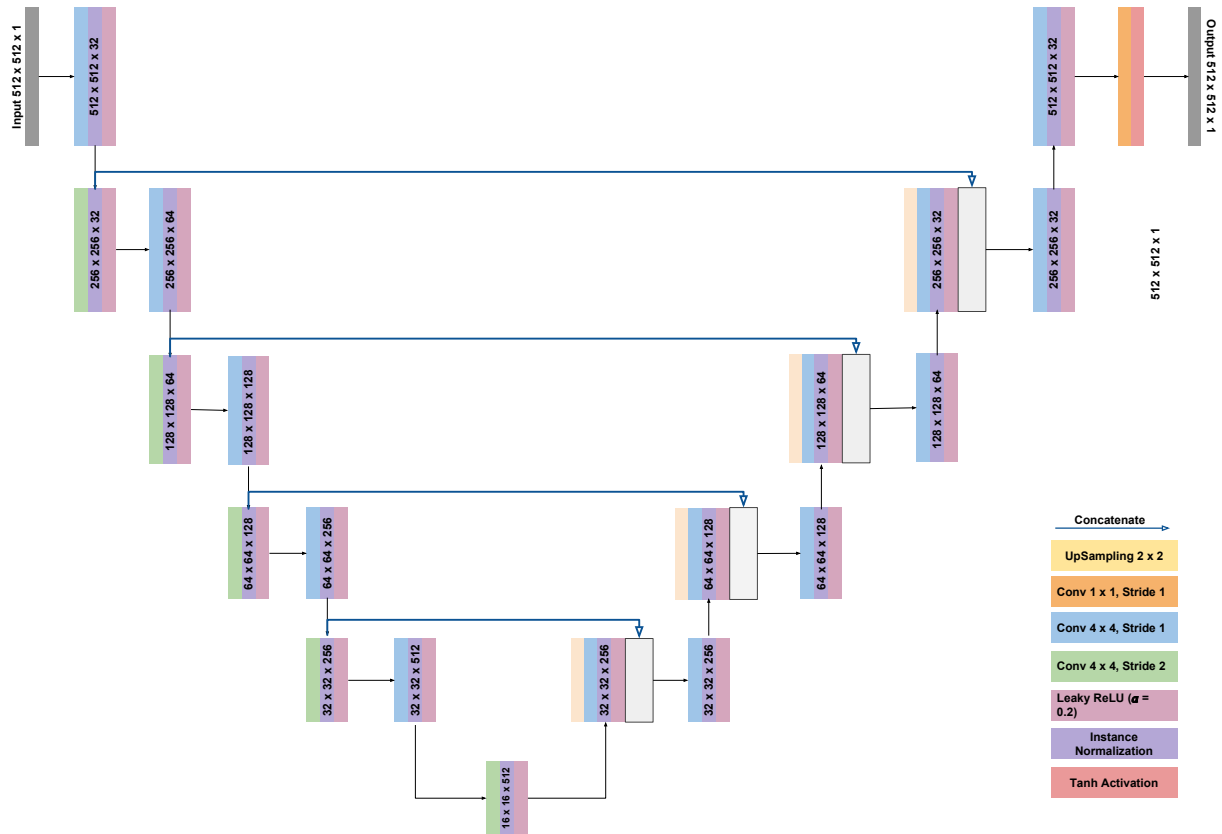


Figure 3.3: U-net

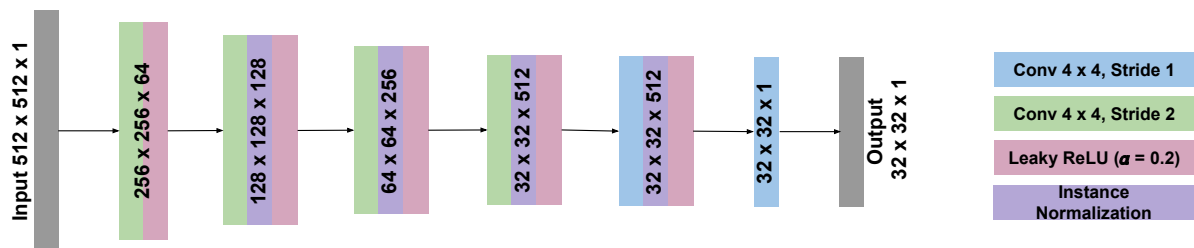


Figure 3.4: CycleGAN discriminator architecture

feature maps for discriminating the synthesized and original training images. All layers of these networks utilize instance normalization and LeakyReLU (rate = 0.2) activation functions, except that the last layers of the generators and discriminators use ‘tanh’ and linear activation functions, respectively.

3.2 Deep Segmentation via Adversarial Learning ³

Adversarial domain adaptation module feeds sCBCT images into the deep segmentation module. Specifically, the deep segmentation network can be pre-trained using sCBCT images with their inherited segmentation masks of original CT images. However, due to the domain shift between CT and CBCT, the CT segmentation masks may not work properly for training the segmentation network in CBCT domain. Thus, the segmentation module is designed to deploy both sCBCT and CBCT images as inputs, creating the corresponding output segmentation maps. Then, a feature (segmentation) discriminator D_{feat} is integrated to determine whether the segmentation maps are from CT or CBCT domains (Fig. 3.1). The segmentation networks can be considered as the corresponding generators of segmentation maps in GAN. When two generators can produce segmentation maps to fool D_{feat} so that the corresponding sCBCT segmentation maps and CBCT maps are indistinguishable, we achieve a good segmentation network for CBCT images. Through this operation, the segmentation network can learn more info about target domain (CBCT) via adversarial learning, and predict on target images more reliably even without any label.

The segmentation network is denoted as Seg . To overcome the potential unbalance in medical image segmentation masks, the loss function is chosen to be based on the DSC (DICE Similarity Coefficient) [48]:

$$L_{seg} = 1 - \frac{2Y_{CT}Y_{map} + a}{Y_{CT} + Y_{map} + a}, \quad (3.6)$$

where Y_{CT} is the CT segmentation mask (Ground Truth), Y_{map} is the output of the deep model. The smoothing term a ensures the stability of the loss function by avoiding potential numerical issues when the denominator becomes 0.

As shown in Fig. 3.1, D_{feat} aims to distinguish the segmentation maps of two domains. The loss function of D_{feat} is given as:

$$L_{adv} = E_{x \sim \mathcal{P}_{CBCT}(x)}[D_{feat}(Seg(x))^2] + E_{z \sim \mathcal{P}_{sCBCT}(z)}[(1 - D_{feat}(Seg(z)))^2], \quad (3.7)$$

³Part of this section is reprinted from our accepted manuscript [47] ©2019 MICCAI

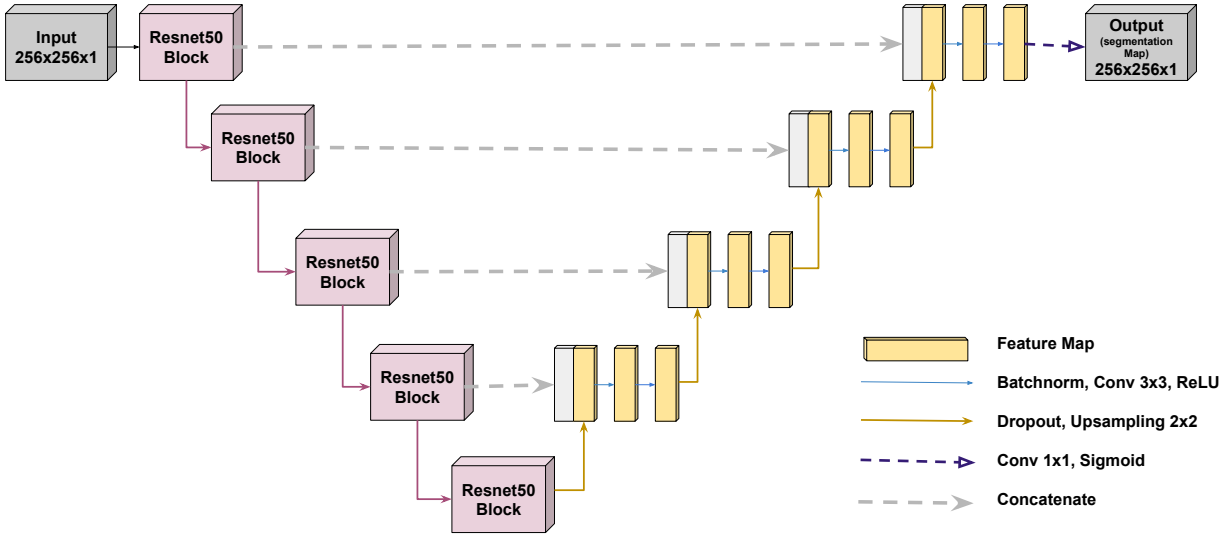


Figure 3.5: Segmentation module

where P_{CBCT} , P_{sCBCT} are the sets of input CBCT and sCBCT images. Finally, we establish the total loss function:

$$L_{feat} = L_{adv} + \lambda_{seg} L_{seg}. \quad (3.8)$$

In the segmentation module, the Resnet50 [49] block is built in the encoder part of U-net [28] as the basic architecture. The architecture as shown in 3.5.

Again, we choose PatchGAN [34] as a discriminator with the receptive field size 70×70 . As shown in 3.6, the network consists of 5 convolutional layers with 4×4 kernel size and stride of 2, except for the last one layers with convolution stride of 1. The numbers of feature maps are 64, 128, 256, 512, 1 for each layer, respectively. For the first four layers, each convolutional layer is followed by a leaky ReLU with rate = 0.2 and an instance normalization layer. Through the integration of adversarial domain adaptation and segmentation modules, the available CT images with segmentation masks as well as CBCT images are taken the best advantage of when training for CBCT segmentation. In the training, we first train D_{feat} until the accuracy reaches an accuracy threshold \mathcal{R} (initial setting is 0.6). Next, we train the segmentation part until we get a reliable

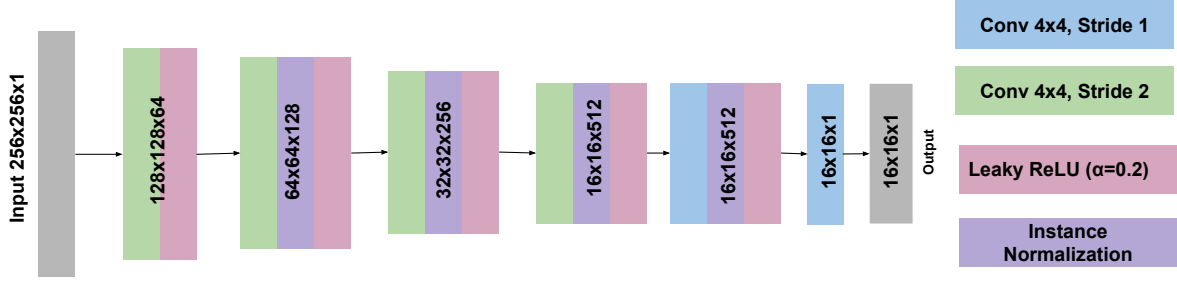


Figure 3.6: Feature discriminator architecture

module for CBCT.

3.3 Learning Diagram

Importantly, a key characteristic in our proposed method is to do domain adaptation in both image and feature levels. More specifically, CycleGAN is optimized by cycle consistency loss L_{cycle} and L_{GAN} via the image adaptation perspective. The segmentation also collects gradients back-propagated from the discriminators D_{feat} towards feature adaptation. In these regards, the feature discriminator is fitted in a multi-task learning setting, such that, it is able to distinguish CT and CBCT predicted contours and feedback the loss into the segmentation module. In turn, our proposed method has the great improvements, and emphasize pixel-wise cyclic reconstruction and focus on structural semantics.

During training, for the CycleGAN module, its sub-modules are sequentially updated in the following order: $D_{CBCT} \rightarrow D_{CT} \rightarrow G_{CBCT} \rightarrow G_{CT}$. Specifically, the generator G_{CT} and G_{CBCT} are updated first to obtain the fake (target-like) images. Then the discriminator D_{CT} and D_{CBCT} are updated to differentiate the fake images from the actual target images. For the segmentation module, followed by the segmentation part and feature discriminator to map the extracted features to the segmentation predictions and generated target-like images. We first train the feature discriminator, when the accuracy reaches the preset threshold, then update the L_{adv} and L_{seg} and propagate back to the segmentation module for the feature level adaptation. Thus, the final loss function of

our proposed method is as follows:

$$\begin{aligned} L = & L_{GAN}(G_{CT}, D_{CT}) + L_{GAN}(G_{CBCT}, D_{CBCT}) \\ & + \lambda_{cycle} L_{cycle}(G_{CT}, G_{CBCT}) + \lambda_{id} L_{identity}(G_{CT}, G_{CBCT}) + L_{adv} + \lambda_{seg} L_{seg} \end{aligned} \quad (3.9)$$

4. EXPERIMENTS AND RESULTS

4.1 Datasets ¹

The data is collected 90 patients' CBCT and CT images from UT Southwestern Medical School, Department of Radiation Oncology. For each patient, we select one planning CT and one CBCT to perform the CycleGAN training. As shown in Fig. 4.1 and Fig. 4.2, we can visualize that CBCT images have more scatters, artifacts, and sometimes truncated regions than CT images do.

In our experiments, the pixel spacing of CT and CBCT are normalized into 1×1 , where the thickness is 3mm. All slices are cropped into 512×512 resolution. Each CBCT study has 88 slices, and each CT study has about 220 slices, while we crop CT into 88 slices which focus on bladder and prostate portions. The HU value range for CT is $[-1000, 3500]$, and that for CBCT is $[-1000, 7000]$. All the image HU values are normalized to $(-1, 1)$ for training and validation. We use the manual segmentation of CBCT slices as the ground truth to validate the proposed method in this thesis.

4.2 Implementation Details ²

The networks are implemented and trained using Keras [50] on a PC with an NVIDIA Tesla K80 dual-GPU. For CycleGAN [8], the following hyper-parameters are set for training the CycleGAN module: the batch size is set to 1 with ADAM [51] for optimization at the learning rate 0.003 with scheduled decay rate at 0.005; $\beta_1 = 0.9$ and the epoch number = 100. For the CycleGAN loss, we set $\lambda_{cycle} = 10$, $\lambda_{id} = 5$.

For segmentation, we first train U-net with the Resnet50 backbones on both CT and sCBCT datasets. Owing to data imbalance and memory limits, we crop all the images into 256×256 . The following hyper-parameters are adopted for training: setting batch size as 20 and using the ADAM optimizer with learning rate at 0.03 $\beta_1 = 0.9$ and epoch number = 300.

¹Part of this section is reprinted from our accepted manuscript [47] ©2019 MICCAI

²Part of this section is reprinted from our accepted manuscript [47] ©2019 MICCAI

For feature-level adversarial learning in the segmentation module, We still use ADAM as the optimizer with learning rate 0.03. Due to memory limit, the batch size is set to 2, epoch number is set to 200. To get the best bladder segmentation performance, we have explored different loss weight λ_{seg} settings during training.

4.3 Experimental Results & Discussion ³

We use 90 patients’ data, including CBCT images and planning CT images with outlined contours, to train our model—Unsupervised segmentation based on Feature and Pixel Domain Adaptation (UFPDA), specifically its constituting CycleGAN, supervised segmentation, and feature discriminator modules. Fig. 4.3 visualizes image domain adaptation via CycleGAN. To better visualize the noise and scatters, we have set the range of displayed intensity to [-200,400] because all soft tissue and main organ’s HU values are observed to be concentrated in this range. From Fig 4.3, we can see that the sCBCT images from our domain adaptation generator have higher noise and lower contrast compared with the original CT images; while it keeps the CT anatomical structures when compared with the corresponding CBCT image appearance.

To evaluate our CBCT segmentation algorithm using adversary learning domain adaption, we compare the derived segmentation with our manually labeled segmentations of 676 hold-out slices from 7 patients. We emphasize here again that one of critical challenges for CBCT image analysis using machine learning is that it has been very difficult to collect a large enough set of high-quality manual CBCT segmentations for training due to both poor image quality and required intensive labor. This is the exact motivation of this thesis. We have evaluated our model on this subset of 676 CBCT slices (note that we never used the CBCT segmentation in training). All the CBCT slices from the 7 patients have been manually labeled by radiation physicians.

To evaluate the improvement after the domain adaptation module in our method, we adopt the DICE similarity coefficient (DSC) as the evaluation metric [52]. The DSC score is defined as:

$$DSC = \frac{2(Y_{GT} \cap Y_{map})}{Y_{GT} \cup Y_{map}}, \quad (4.1)$$

³Part of this section is reprinted from our accepted manuscript [47] ©2019 MICCAI

Method	Pixel	Feature	DSC (%)
Source only	no	no	70.1
CycleGAN	yes	no	75.8
UFPDA (Ours)	yes	yes	83.6

Table 4.1: Performance comparison for CBCT bladder segmentation, ‘Pixel’ means use CycleGAN to transfer CT image domain into CBCT, ‘Feature’ means use feature discriminator to do adversarial learning in segmentation module.

where Y_{GT} is the image segmentation mask (Ground Truth), Y_{map} is the output of the deep model. DSC is similar to the Jaccard index as a commonly adopted segmentation evaluation metric. Its range is $[0, 1]$ based on the ratio of the intersection to the union of the predicted segmentation and ground-truth mask. We will show the DSC from different segmentation methods in the CBCT validation dataset. Fig. 4.4 shows the segmentation maps by different methods together with the ground truth. In Fig. 4.4, we can visually observe the improvement by our approach integrating cycle-consistent domain adaption and adversary learning for segmentation. Specifically, ‘Source only’ represents the results by segmenting CBCT images directly using the CT pre-training segmentation model without any domain adaptation. ‘CycleGAN’ denotes the results by using CycleGAN for pixel-level domain adaptation to generate sCBCT images and then train the segmentation network using only sCBCT images and inherited CT segmentation masks. ‘UFPDA’ (Unsupervised segmentation based on Feature and Pixel Domain Adaption) is our proposed method, combining the pixel-level domain adaptation and feature-level adversarial learning. Table 4.2 shows the comparison of sensitivity analysis. With $\lambda_{seg} = 2$, we achieve the best segmentation performance with DSC 83.6%.

Table 4.1 shows the DSC of ablation studies. Table 4.2 shows the comparison of sensitivity analysis with λ_{seg} . With $\lambda_{seg} = 2$, we achieve the best segmentation performance with DSC 83.6%. As shown in Fig. 4.4 and Table 4.1, our proposed method (UFPDA) improves CBCT segmentation performance by about 7.8% for the DSC score, compared to the naive implementation of CycleGAN to synthesize CBCT images.

Based on the method we proposed, and the results we have, we found that, even though CBCT

λ_{seg}	1	2	3	4	5	6	7	8	9	10
DSC (%)	82.3	83.6	80.3	80.8	79.3	80.1	80.7	78.9	80.8	79.9

Table 4.2: Sensitivity analysis from λ_{seg} 1 to 10

has inaccurate Hounsfield Unit (HU) values and more artifacts compared with CT, our developed deeper generator is able to generate more faithful 512 x 512 synthesised CBCT (sCBCT) to match with the CBCT image quality. To handle transformation distortion, the adversarial loss, cycle-consistency loss and identity loss help prevent the generators from having synthetic images with distortion irrelevant to the input. The identity loss contributes to regularize the generator to be near an identity mapping when something already looks like from the target domain. For example, the identity loss helps preserve the HU values of CT when we try to translate CT to sCBCT. We also have adopted patchGAN as the discriminator, together with the MSE loss functions to correctly discriminate the predicted sCBCT and CBCT segmentation maps, with an adversarial loss feeding back to the segmentation model with the needed feature level adaptation. The difficulty in obtaining CBCT segmentation labels and the inferior image quality of CBCT yield a unique data challenge. And, for unbalanced source and target data, we first normalize all CT and CBCT into same slice thickness (3mm), then use all 88 CBCT slices and randomly select 88 slices from CT for training. Our model adapts planning CT’s segmentation labels to the CBCT domain for finer segmentation and has great clinical impact in adaptive radiotherapy. Patients with large anatomy changes during radiation treatment can benefit from accurate contours for adaptive planning with minimum human intervention. It will lift the major obstacle in adopting adaptive radiation therapy.

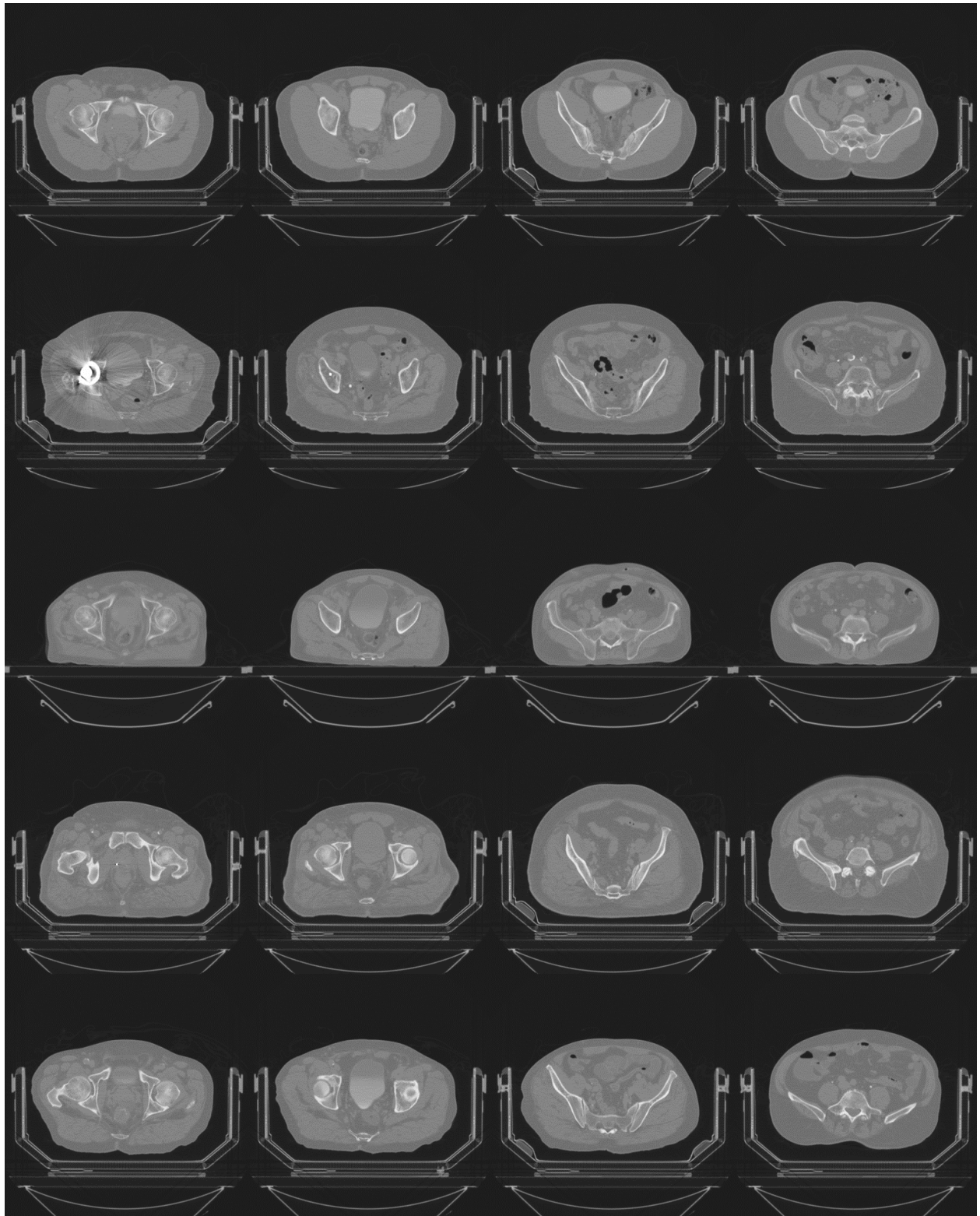


Figure 4.1: Visualization of example CT images

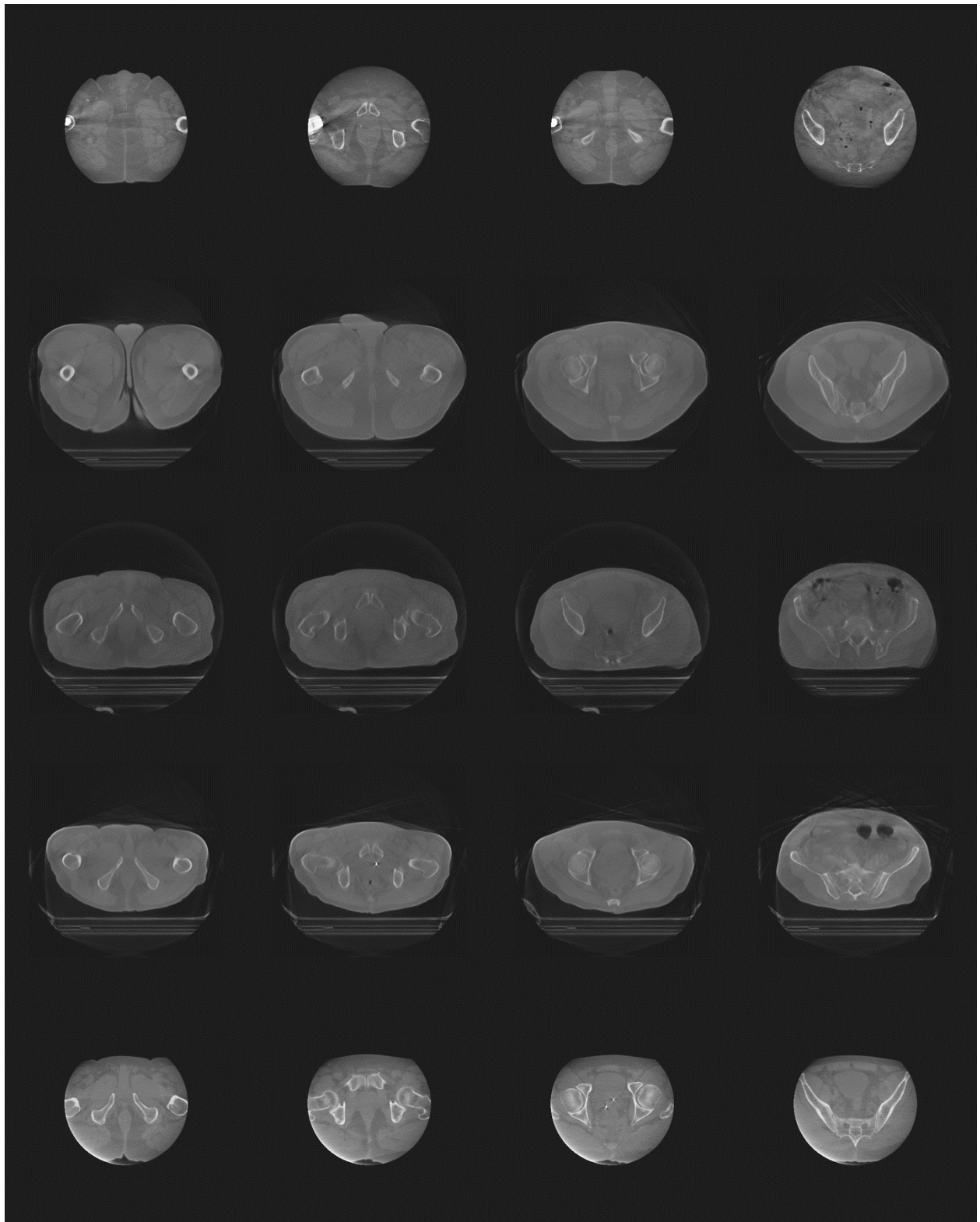


Figure 4.2: Visualization of example CBCT images

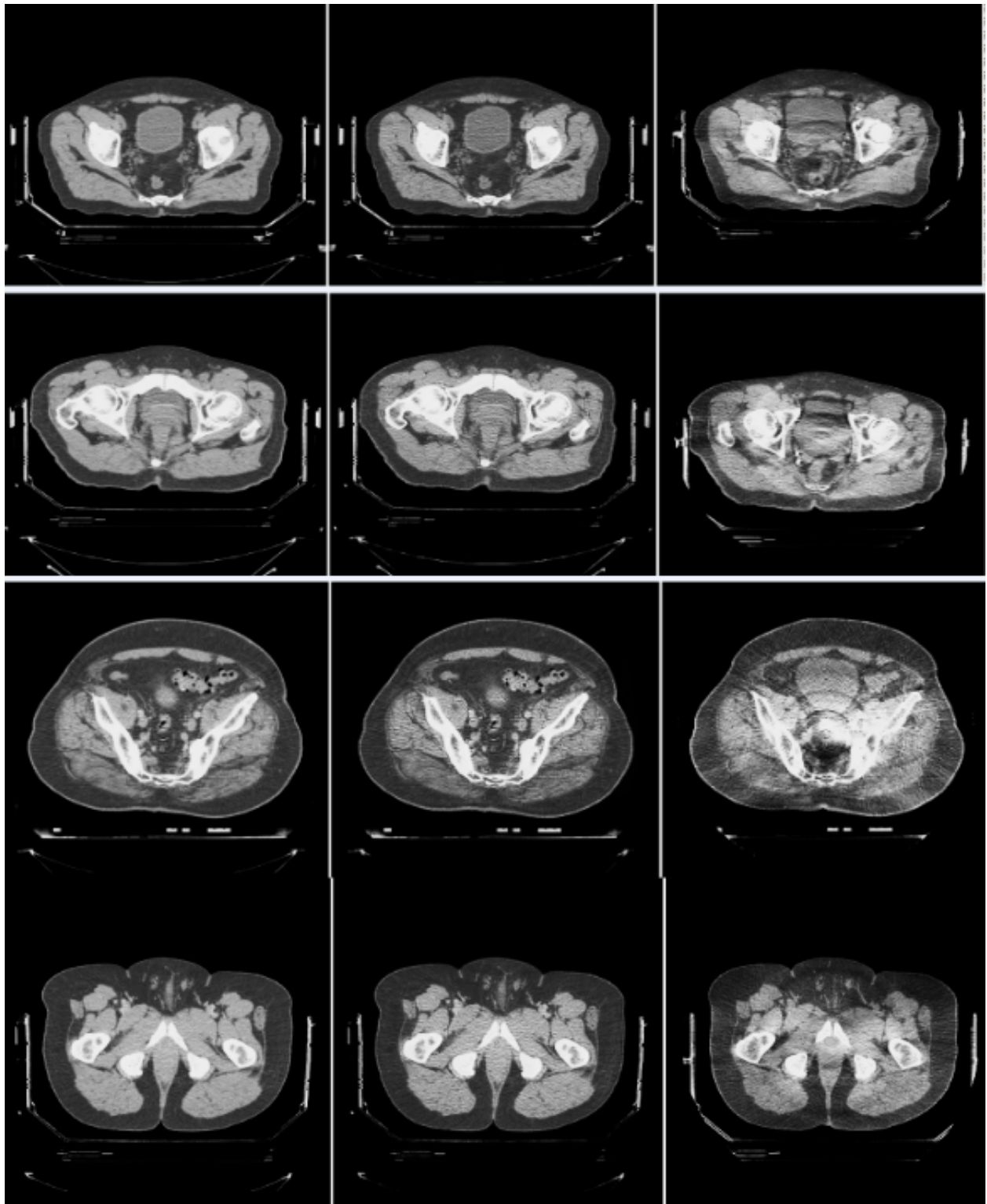


Figure 4.3: Domain adaptation between CT and CBCT: From left to right are displayed raw CT, sCBCT, and CBCT images.

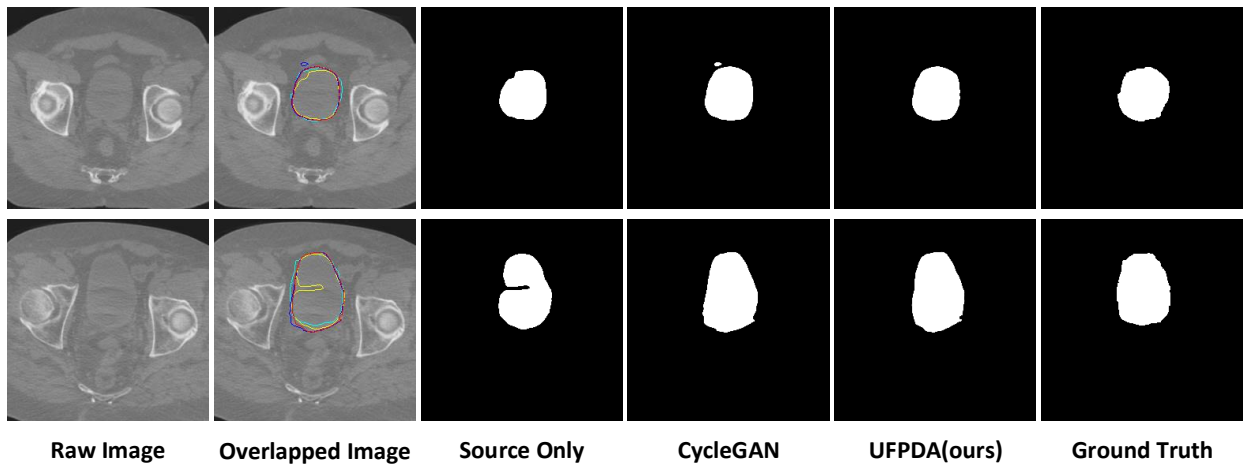


Figure 4.4: Visual comparison of segmentation results by different methods. From left to right: raw CBCT image, Overlapped Image, segmentation by ‘Source only’ (yellow in Overlapped Image), ‘CycleGAN’ (blue) and UFPDA (red), Ground Truth (light blue).

5. CONCLUSION AND FUTURE WORK

In this thesis, we have proposed a method combining the cycle-consistent domain adaptation and adversarial learning for CBCT image segmentation. This is the first time CycleGAN and CyCADA are integrated to address CBCT segmentation with domain adaptation. Although motivated by CycleGAN and the domain adaptation scheme is similar to CyCADA, we have to refine the architecture and several components for the segmentation-oriented domain adaptation between CBCT and CT images, for which image features are significantly different. The method can generate reasonable CBCT segmentation without the need for labelled CBCT segmentation for training. In our studies using 90 clinical bladder CBCT datasets, our method enables training with both CT and CBCT images and improves the DSC score by about 13.6 percent as compared with the segmentation model trained only with CT images. The proposed method can be extended to other applications where existing segmentation labels can be transferred to the datasets in new domains.

Even though our method does improve CBCT segmentation accuracy, There are some unresolved questions:

1) The modules are still training stage by stage; that is, we first train the image domain adaptation module, then train image segmentation. The segmentation loss or feature adversarial loss is not fed back into the generative module. Thus, for our future work, we intend to design an end-to-end architecture for domain adaptation and segmentation. The feature adversarial loss and segmentation loss will help to generate better sCBCT, which will improve the segmentation performance consequently.

2) Currently, we only use the segmentation map as the feature discriminator input. However, the original sCBCT and CBCT images may contain more information to train the discriminator. It might be better to expand the input of two channels: one channel for the image, the other channel for the segmentation map. The new two-channel discriminator may be more capable of distinguishing CBCT and sCBCT images, and therefore improve the overall performance.

3) The medical images, like MRI, CT, CBCT, are often 3D datasets. Information in the z-axis

was not used in our project. Balagopal *et al.* [10] designed a 3D segmentation module, and got the best performance for prostate CT segmentation to the best of our knowledge. The authors in [53] proposed a 3D version Autoencoder architecture for segmentation. To utilize 3D information for CBCT, it would be better to design 3D modules for domain adaptation and segmentation.

We can re-design the image generator as a simpler network structure, but use 3-dimensional convolutional network, for example 3D-AutoEncoder, to design a 3D version segmentation network and feature discriminator. The final end-to-end network will consist of these three parts as illustrated in the figure. Such a design presents a simpler network than the combination of CycleGAN and U-net segmentation, which enables scalable learning with 3D images considering both required memory and computational resources. By better leveraging 3D datasets, this model may get more accurate CBCT segmentation compared to using 2D data to further improve adaptive radiation therapy development.

REFERENCES

- [1] D. Yan, F. A. Vicini, J. D. Wong, and A. A. P. Martinez, “Adaptive radiation therapy,” *Physics in medicine and biology*, vol. 42 1, pp. 123–32, 1997.
- [2] S. Oh and S. Kim, “Deformable image registration in Radiation Therapy,” in *Radiation oncology*, 2017.
- [3] R. W. Schulze, U. Heil, D. D. Groß, D. Brüllmann, E. Dranischnikow, U. Schwanecke, and E. Schömer, “Artifacts in CBCT: a review,” *Dento maxillo facial radiology*, vol. 40 5, 2011.
- [4] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, “Generative adversarial nets,” in *Conference on Neural Information Processing Systems*, 2014.
- [5] D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes,” *ArXiv*, vol. abs/1312.6114, 2013.
- [6] H. Yang *et al.*, “Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN,” *ArXiv*, 2018.
- [7] X. Liang, L. Chen, D. Nguyen, Z. Zhou, X. Gu, M. Yang, J. Wang, and S. B. Jiang, “Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy.,” *Physics in medicine and biology*, 2018.
- [8] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *International Conference on Computer Vision*, 2017.
- [9] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Conference on Computer Vision and Pattern Recognition*, 2015.

- [10] A. Balagopal, S. Kazemifar, D. Nguyen, M. H. Lin, R. Hannan, A. Owrangi, and S. Jiang, “Fully automated organ segmentation in male pelvic CT images,” *Physics in medicine and biology*, vol. 63 24, p. 245015, 2018.
- [11] Q. Dou, C. Ouyang, C. Chen, H. Chen, and P. A. Heng, “Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss,” *ArXiv*, vol. abs/1804.10916, 2018.
- [12] C. Chen, Q. Dou, H. Chen, J. Qin, and P. A. Heng, “Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation,” *ArXiv*, vol. abs/1901.08211, 2019.
- [13] P. Hüttenrauch, M. L. I. Witt, D. Wolff, S. Bosold, R. Engenhardt-Cabillic, J. A. Sparenberg, H. Vorwerk, and K. Zink, “Target volume coverage and dose to organs at risk in prostate cancer patients,” *Strahlentherapie und Onkologie*, vol. 190, pp. 310–316, 2013.
- [14] J. Hofmaier, J. Haehnle, C. Kurz, G. Landry, C. Maihoefer, L. H. Schüttrumpf, P. Süß, K. Teichert, M. Soehn, N. Spahr, C. Brachmann, F. Weiler, C. Thieke, K.-H. Küfer, C. Belka, K. Parodi, and F. Kamp, “Multi-criterial patient positioning based on dose recalculation on scatter-corrected CBCT images.,” *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 125 3, pp. 464–469, 2017.
- [15] T. Li, D. Thongphiew, X. Zhu, W. R. Lee, Z. Vujasković, F.-F. Yin, and Q. J. Wu, “Adaptive prostate IGRT combining online re-optimization and re-positioning: a feasibility study.,” *Physics in medicine and biology*, vol. 56 5, pp. 1243–58, 2011.
- [16] D. Thongphiew, Q. J. Wu, W. R. Lee, V. Chankong, S. Yoo, R. L. McMahon, and F.-F. Yin, “Comparison of online IGRT techniques for prostate IMRT treatment: adaptive vs repositioning correction.,” *Medical physics*, vol. 36 5, pp. 1651–62, 2009.
- [17] Q. J. Wu, D. Thongphiew, Z. Wang, B. Mathayomchan, V. Chankong, S. Yoo, W. R. Lee, and F.-F. Yin, “On-line re-optimization of prostate IMRT plans for adaptive radiation therapy.,” *Physics in medicine and biology*, vol. 53 3, pp. 673–91, 2008.

- [18] A. Qin, Y. Sun, J. Z. Liang, and D. i Yan, “Evaluation of online/offline image guidance/adaptation approaches for prostate cancer radiation therapy,” *International journal of radiation oncology, biology, physics*, vol. 91 5, pp. 1026–33, 2015.
- [19] W. Crijs, H. V. Herck, G. Defraene, L. V. den Bergh, P. Slagmolen, K. Haustermans, F. Maes, and F. van den Heuvel, “Dosimetric adaptive IMRT driven by fiducial points.,” *Medical physics*, vol. 41 6, p. 061716, 2014.
- [20] R. Boggula, F. Lorenz, Y. Abo-Madyan, F. Löhr, D. Wolff, J. Boda-Heggemann, J. W. Hesser, F. Wenz, and H. Wertz, “A new strategy for online adaptive prostate radiotherapy based on cone-beam CT.,” *Zeitschrift fur medizinische Physik*, vol. 19 4, pp. 264–76, 2009.
- [21] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” in *Proceedings of the IEEE*, pp. 2278–2324, 1998.
- [22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Conference on Neural Information Processing Systems*, 2012.
- [23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *Computer Vision and Pattern Recognition*, 2015.
- [24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” *Computer Vision and Pattern Recognition*, 2015.
- [25] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-V4, Inception-ResNet and the impact of residual connections on learning,” in *ICLR 2016 Workshop*, 2016.
- [26] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Computer Vision and Pattern Recognition*, 2015.
- [27] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 2481–2495, 2015.

- [28] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *MICCAI*, 2015.
- [29] X. Li, J. Zhou, X. Qi, Q. Dou, C.-W. Fu, and P. A. Heng, “H-DenseUNet: Hybrid densely connected U-Net for liver and tumor segmentation from CT volumes,” *IEEE Transactions on Medical Imaging*, vol. 37, pp. 2663–2674, 2017.
- [30] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “UNet++: A nested U-Net architecture for medical image segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018.
- [31] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *CoRR*, vol. abs/1511.06434, 2015.
- [32] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive growing of GANs for improved quality, stability, and variation,” *ArXiv*, vol. abs/1710.10196, 2017.
- [33] A. Brock, J. Donahue, and K. Simonyan, “Large scale GAN training for high fidelity natural image synthesis,” *ArXiv*, vol. abs/1809.11096, 2018.
- [34] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Conference on Computer Vision and Pattern Recognition*, 2017.
- [35] M. Long, Y. Cao, J. Wang, and M. I. Jordan, “Learning transferable features with deep adaptation networks,” in *International Conference on Machine Learning*, 2015.
- [36] F. Zhuang, X. Cheng, P. Luo, S. J. Pan, and Q. He, “Supervised representation learning: Transfer learning with deep Auto-Encoders,” in *IJCAI*, 2015.
- [37] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, “Adversarial discriminative domain adaptation,” *Computer Vision and Pattern Recognition*, 2017.
- [38] Z. Cao, M. Long, J. Wang, and M. I. Jordan, “Partial transfer learning with selective adversarial networks,” *Computer Vision and Pattern Recognition*, 2017.

- [39] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” in *International Conference on Machine Learning*, 2017.
- [40] J. Shen, Y. Qu, W. Zhang, and Y. Yu, “Wasserstein distance guided representation learning for domain adaptation,” in *AAAI*, 2017.
- [41] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. A. Efros, and T. Darrell, “CyCADA: Cycle consistent adversarial domain adaptation,” in *International Conference on Machine Learning*, 2018.
- [42] F. Milletari, N. Navab, and S. A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” *2016 Fourth International Conference on 3D Vision*, 2016.
- [43] M. Ran, J. Hu, Y. Chen, H. Chen, H. Sun, J. Zhou, and Y. Zhang, “Denoising of 3D magnetic resonance images using a residual Encoder-Decoder Wasserstein generative adversarial network,” *Medical image analysis*, vol. 55, pp. 165–180, 2019.
- [44] I. Öksüz, J. R. Clough, A. Bustin, G. Cruz, C. Prieto, R. M. Botnar, D. Rueckert, J. A. Schnabel, and A. P. King, “Cardiac MR motion artefact correction from K-space using deep Learning-Based reconstruction,” in *MLMIR@MICCAI*, 2018.
- [45] C. Chen, Q. Dou, J. Zhou, and P. A. Heng, “Semantic-Aware generative adversarial nets for unsupervised domain adaptation in chest X-Ray segmentation,” *ArXiv*, vol. abs/1806.00600, 2018.
- [46] Q. Dou, C. Ouyang, C. Chen, H. Chen, B. Glocker, X. Zhuang, and P. A. Heng, “PnP-AdaNet: Plug-and-play adversarial domain adaptation network with a benchmark at cross-modality cardiac segmentation,” *ArXiv*, vol. abs/1812.07907, 2018.
- [47] X. Jia, S. Wang, X. Liang, A. Balagopal, D. Nguyen, M. Yang, Z. Wang, J. X. Ji, X. Qian, and S. Jiang, “Cone-Beam computed tomography (CBCT) segmentation by adversarial learning domain adaptation,” in *MICCAI*, 2019.

- [48] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, “Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations,” *ArXiv*, vol. abs/1707.03237, 2017.
- [49] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Conference on Computer Vision and Pattern Recognition*, 2016.
- [50] F. Chollet *et al.*, “Keras.” <https://keras.io>, 2015.
- [51] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *CoRR*, vol. abs/1412.6980, 2014.
- [52] R. R. Shamir, Y. Duchin, J. Kim, G. Sapiro, and N. Y. Harel, “Continuous dice coefficient: a method for evaluating probabilistic segmentations,” *ArXiv*, vol. abs/1906.11031, 2018.
- [53] A. Myronenko, “3D MRI brain tumor segmentation using Auto-Encoder regularization,” in *BrainLes@MICCAI*, 2018.