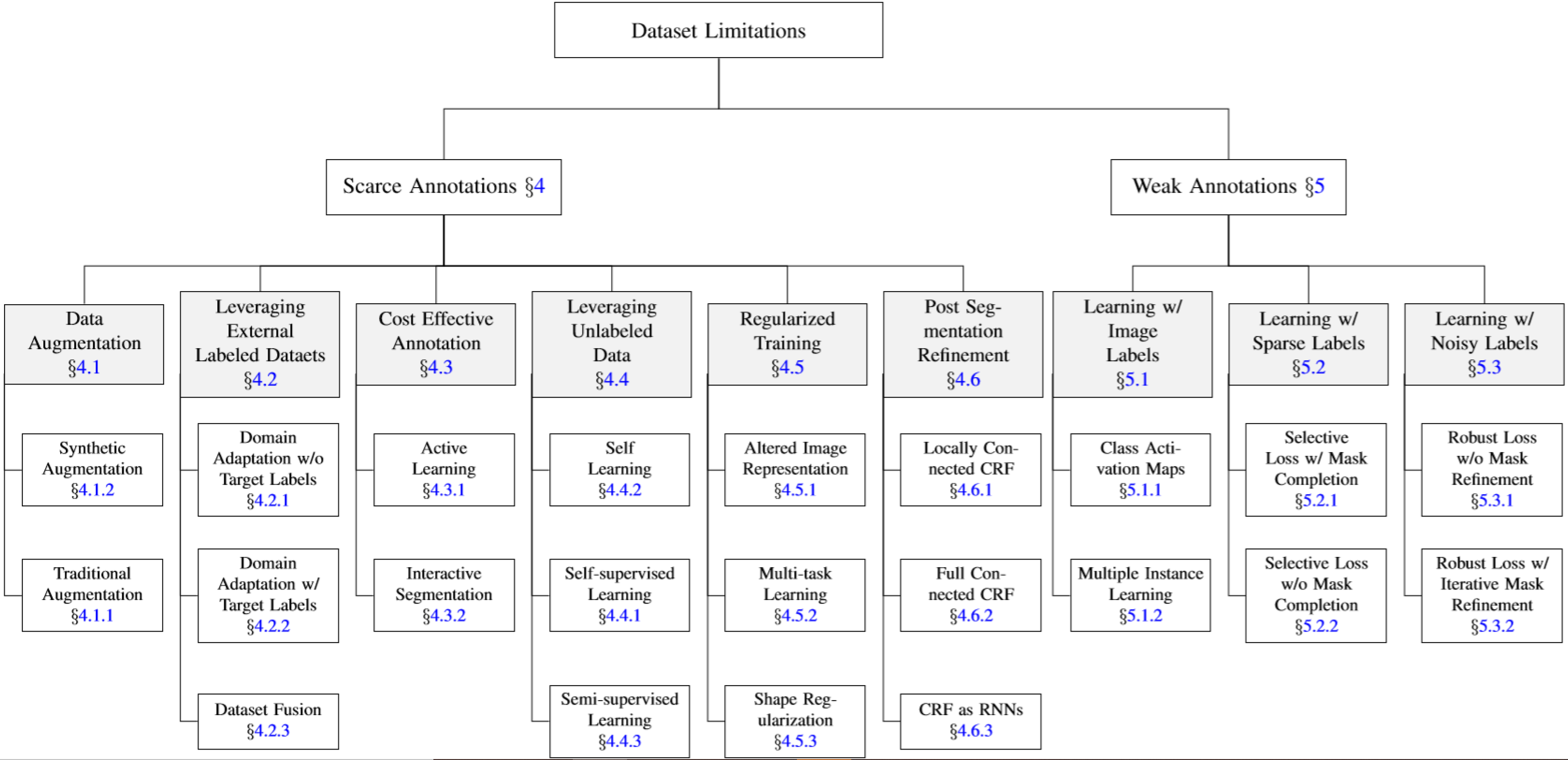
Embracing Imperfect Datasets: A Review of Deep Learning Solutions for Medical Image Segmentation

Abstract

- 医学影像学的卷积神经网络模型需要大型、有代表性和高质量的带标注的数据集。
- 然而，很少有一个完美的训练数据集，尤其是在医学成像领域，包括数据和标注。
- 医学图像分割问题两个主要的数据集限制:①稀少的标注，其中只有有限的标注数据可用于训练;②弱标注，其中训练数据只有稀疏标注、噪声标注或图像级标注。



Dataset Limitations
Scarce Annotations §4
Weak Annotations §5
Data Augmentation §4.1
Synthetic Augmentation §4.1.2
Traditional Augmentation §4.1.1
Leveraging External Labeled Dataets §4.2
Domain Adaptation w/o Target Labels §4.2.1
Domain Adaptation w/ Target Labels §4.2.2
Dataset Fusion §4.2.3
Cost Effective Annotation §4.3
Active Learning §4.3.1
Interactive Segmentation §4.3.2
Leveraging Unlabeled Data §4.4
Self Learning §4.4.2
Self-supervised Learning §4.4.1
Semi-supervised Learning §4.4.3
Regularized Training §4.5
Altered Image Representation §4.5.1
Multi-task Learning §4.5.2
Shape Regularization §4.5.3
Post Segmentation Refinement §4.6
Locally Connected CRF §4.6.1
Full Connected CRF §4.6.2
CRF as RNNs §4.6.3
Learning w/ Image Labels §5.1
Class Activation Maps §5.1.1
Multiple Instance Learning §5.1.2
Learning w/ Sparse Labels §5.2
Selective Loss w/ Mask Completion §5.2.1
Selective Loss w/o Mask Completion §5.2.2
Learning w/ Noisy Labels §5.3
Robust Loss w/o Mask Refinement §5.3.1
Robust Loss w/ Iterative Mask Refinement §5.3.2

Data augmentation

- 传统的数据增广
  - 图像质量：加噪、锐化、模糊
  - 图像表征：饱和度、亮度、对比度
  - 像素值分布：**B样条变换**、直方图均衡
- 合成的数据增广

合成的数据增广

Table 1: Comparison between image synthesis methods suggested for medical image segmentation.

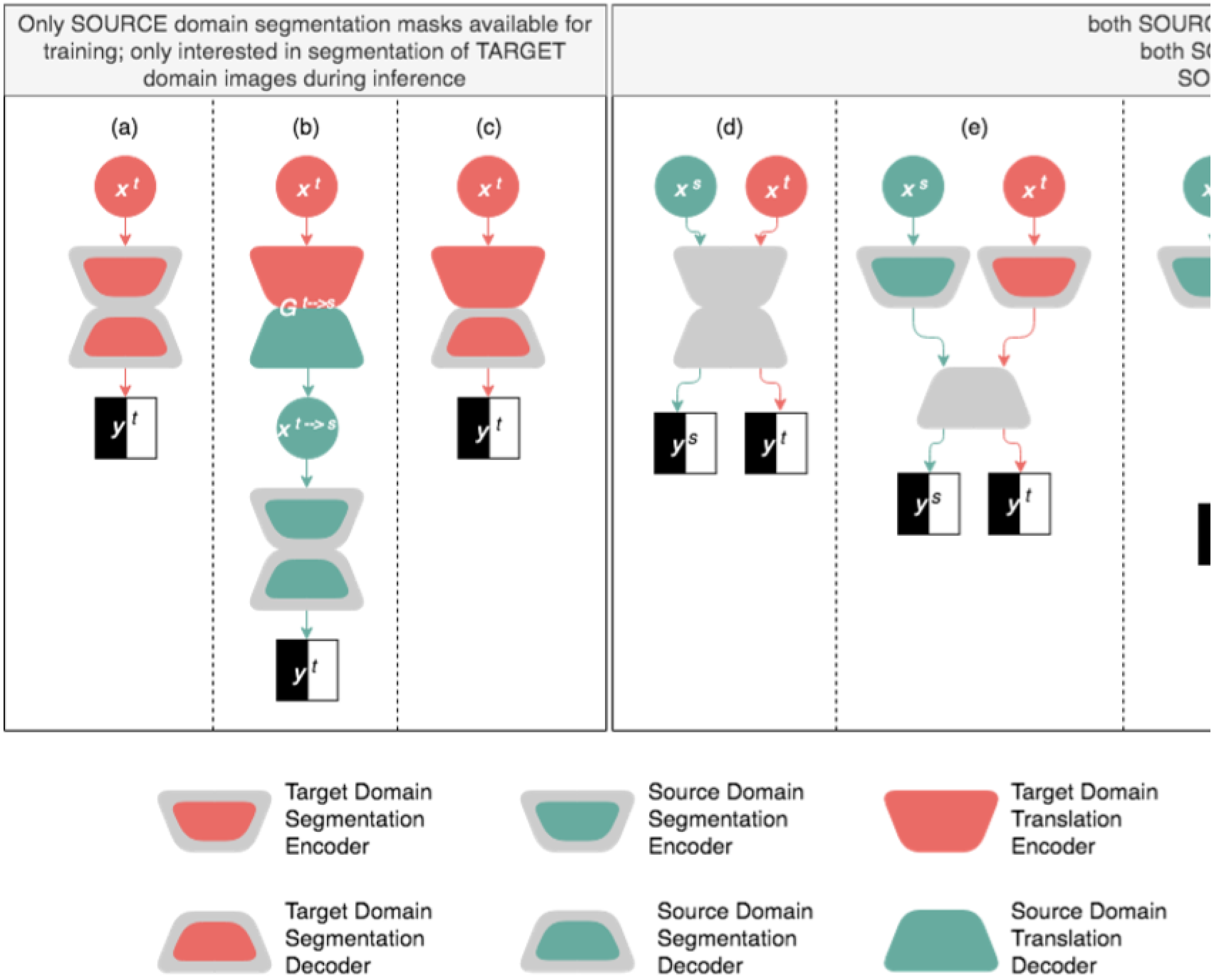
	Synthesis Type	Domains	Description
(2017)	Cross-domain synthesis	CT $\rightarrow$ MRI	Cycle GAN is used to generate pairs of synthesized MR images from pairs of CT slices and the corresponding myocardium masks
18c)	Cross-domain synthesis	CT $\leftrightarrow$MRI	Cycle GAN with shape consistency loss is used to translate between MR and CT scans. Segmentation and synthesis networks are trained jointly.
017)	Same-domain synthesis	Fundus	GAN is used to generate a vessel mask and a conditional GAN is used to generate the corresponding fundus image
8)	Same-domain synthesis	MRI	Conditional GAN to generate synthetic MR images given a lesion mask and a brain segmentation mask
)	Same-domain synthesis	CT	Conditional GAN is used to synthesize pleural nodules in the nodule-free CT slices
9)	Same-domain synthesis	MRI	Hybrid spatial-intensity transformation network is used to synthesize MR images from 1 labeled MR image
. (2018)	Same-domain synthesis	X-ray	Conditional GAN is used to synthesize X-ray images with desired abnormalitie

Leveraging External Labeled Datasets

- 没有标注的域适应
- 带有标注的域适应
- 数据集融合

omain in which segmentation is performed. The Figure column on the right shows the matching data flow from Figure 2.

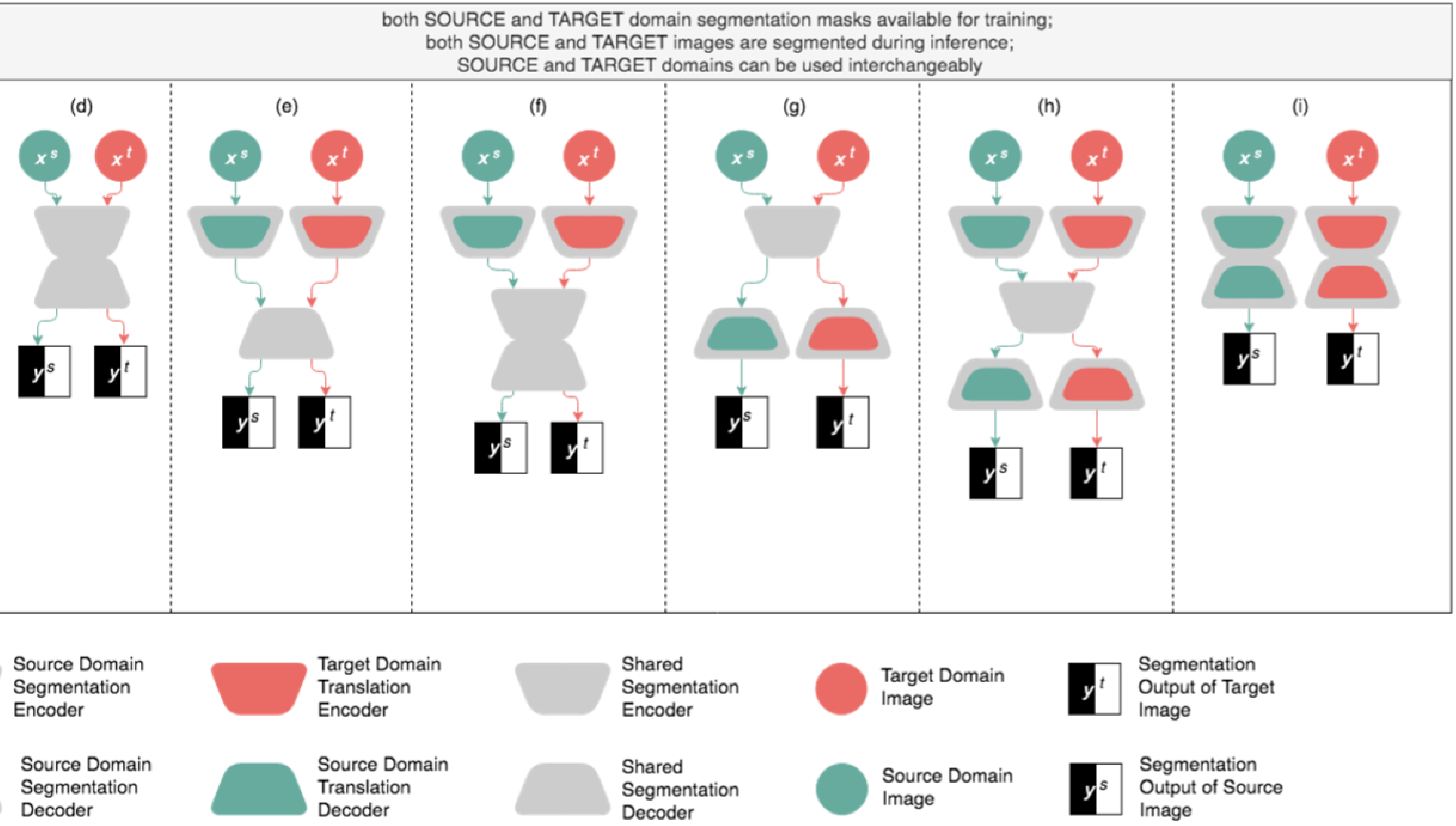
Publication	Availability of Target Domain Segmentation Masks	Segmentation Domain	Modality	Figure
in Adaptation without t Labels				
t al. (2018a)	✖	Target	MRI, CT	(a)
t al. (2018b)	✖	Target	MRI, CT	(a)
et al. (2018)	✖	Source	X-ray	(b)
g et al. (2018b)	✖	Source	DRR, X-ray	(b)
et al. (2019)	✖	Target	MRI, CT	(c)
(2018)	✖	Source	MRI, CT	(b)
in Adaptation with t Labels				
sias et al. (2017)	✔	Both	MRI, CT	(i)
g et al. (2018c)	✔	Both	MRI, CT	(i)
t al. (2018)	✔	Both	MRI, CT	(e)
dria et al. (2018)	✔	Both	MRI, CT	(d),(e),(f),(g),(h)
et Fusion				
ni et al. (2018a)	✔	All domains	MRI,CT,US,X-ray	(d)
iev and Kaufman (2019)	✔	All domains	CT	(d)



Only SOURCE domain segmentation masks available for training; only interested in segmentation of TARGET domain images during inference
both SOURC
both S
SO
(a)
(b)
(c)
(d)
(e)
x^t
y^t
$G^{t\to s}$
$x^{t\to s}$
x^s
y^s
Target Domain Segmentation Encoder
Source Domain Segmentation Encoder
Target Domain Translation Encoder
Target Domain Segmentation Decoder
Source Domain Segmentation Decoder
Source Domain Translation Decoder

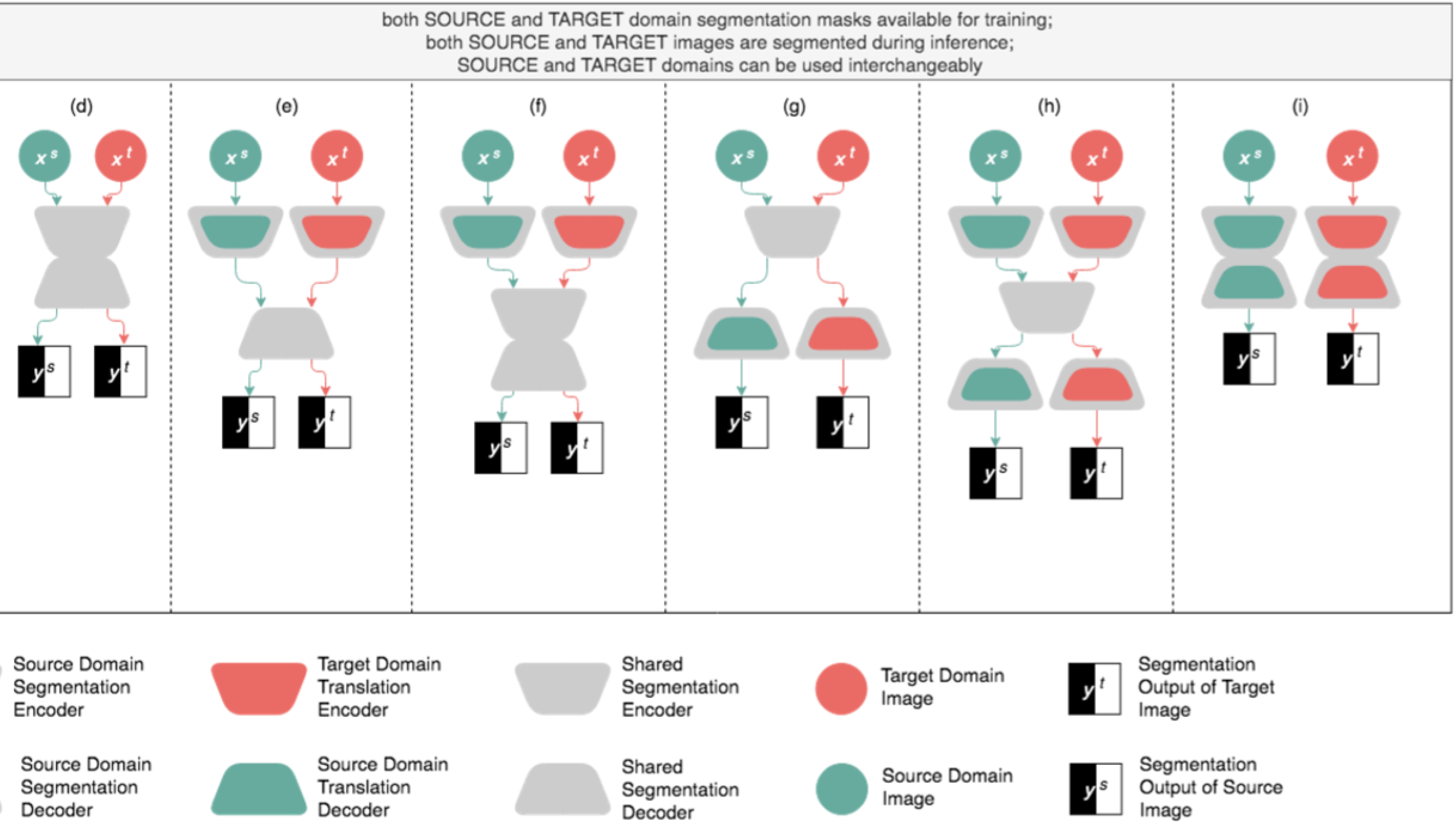
omain in which segmentation is performed. The Figure column on the right shows the matching data flow from Figure 2.

Publication	Availability of Target Domain Segmentation Masks	Segmentation Domain	Modality	Figure
n Adaptation without Labels				
al. (2018a)	✖	Target	MRI, CT	(a)
al. (2018b)	✖	Target	MRI, CT	(a)
t al. (2018)	✖	Source	X-ray	(b)
et al. (2018b)	✖	Source	DRR, X-ray	(b)
t al. (2019)	✖	Target	MRI, CT	(c)
(2018)	✖	Source	MRI, CT	(b)
n Adaptation with Labels				
ias et al. (2017)	✔	Both	MRI, CT	(i)
et al. (2018c)	✔	Both	MRI, CT	(i)
al. (2018)	✔	Both	MRI, CT	(e)
ria et al. (2018)	✔	Both	MRI, CT	(d),(e),(f),(g),(h)
t Fusion				
ni et al. (2018a)	✔	All domains	MRI,CT,US,X-ray	(d)
ev and Kaufman (2019)	✔	All domains	CT	(d)


both SOURCE and TARGET domain segmentation masks available for training;
both SOURCE and TARGET images are segmented during inference;
SOURCE and TARGET domains can be used interchangeably
(d)
(e)
(f)
(g)
(h)
(i)
x^s
x^t
y^s
y^t
Source Domain Segmentation Encoder
Target Domain Translation Encoder
Shared Segmentation Encoder
Target Domain Image
Segmentation Output of Target Image
Source Domain Segmentation Decoder
Source Domain Translation Decoder
Shared Segmentation Decoder
Source Domain Image
Segmentation Output of Source Image

omain in which segmentation is performed. The Figure column on the right shows the matching data flow from Figure 2.

Publication	Availability of Target Domain Segmentation Masks	Segmentation Domain	Modality	Figure
n Adaptation without Labels				
al. (2018a)	✖	Target	MRI, CT	(a)
al. (2018b)	✖	Target	MRI, CT	(a)
t al. (2018)	✖	Source	X-ray	(b)
et al. (2018b)	✖	Source	DRR, X-ray	(b)
t al. (2019)	✖	Target	MRI, CT	(c)
(2018)	✖	Source	MRI, CT	(b)
n Adaptation with Labels				
ias et al. (2017)	✔	Both	MRI, CT	(i)
et al. (2018c)	✔	Both	MRI, CT	(i)
al. (2018)	✔	Both	MRI, CT	(e)
ria et al. (2018)	✔	Both	MRI, CT	(d),(e),(f),(g),(h)
t Fusion				
ni et al. (2018a)	✔	All domains	MRI,CT,US,X-ray	(d)
ev and Kaufman (2019)	✔	All domains	CT	(d)


both SOURCE and TARGET domain segmentation masks available for training;
both SOURCE and TARGET images are segmented during inference;
SOURCE and TARGET domains can be used interchangeably
(d)
(e)
(f)
(g)
(h)
(i)
Source Domain Segmentation Encoder
Target Domain Translation Encoder
Shared Segmentation Encoder
Target Domain Image
Segmentation Output of Target Image
Source Domain Segmentation Decoder
Source Domain Translation Decoder
Shared Segmentation Decoder
Source Domain Image
Segmentation Output of Source Image

提高标注效率

- 主动学习
- 迭代分割

主动学习

Algorithm 1: Active learning

Input : Initial model $\mathcal{M}_0$, unlabeled dataset $\mathcal{U}_0$, size of query batch k, iteration times $\mathcal{T}$, active learning algorithm $\mathcal{A}$

Output: Labeled dataset $\mathcal{L}_\mathcal{T}$, updated model $\mathcal{M}_\mathcal{T}$

$\mathcal{L}_0 \leftarrow \varnothing$;

for $i \leftarrow 1$ **to** $\mathcal{T}$ **do**

  `/* phase 1: query batch selection */`

  $\mathcal{Q}_t \leftarrow \mathcal{A}(\mathcal{U}_{t-1}, \mathcal{M}_{t-1}, k)$;

  annotate samples in $\mathcal{Q}_t$;

  `/* phase 2: update model */`

  $\mathcal{L}_t \leftarrow \mathcal{L}_{t-1} \cup \{(\mathbf{x}, y) | \mathbf{x} \in \mathcal{Q}_t, y \in \mathcal{Y}_t\}$;

  $\mathcal{M}_t \leftarrow$ fine-tuning $\mathcal{M}_{t-1}$ using $\mathcal{L}_t$;

  $\mathcal{U}_t \leftarrow \mathcal{U}_{t-1} \backslash \mathcal{Q}_t$;

end

return $\mathcal{L}_\mathcal{T}, \mathcal{M}_\mathcal{T}$

Publication	Type	Sample selection strategy			Annotation uni
		informativeness	diversity	annotation cost	
Gorriz et al. (2017)	iterative	✓			whole 2D imag
Yang et al. (2017)	iterative	✓	✓		whole 2D imag
Ozdemir et al. (2018)	iterative	✓	✓		whole 2D imag
Kuo et al. (2018)	iterative	✓	✓	✓	whole 3D imag
Sourati et al. (2018)	iterative	✓	✓		2D image patch
Mahapatra et al. (2018)	1-shot	✓			whole 2D imag
Sourati et al. (2019)	iterative	✓	✓		2D image patch
Zheng et al. (2019)	iterative		✓		2D image patch

交互式分割

Algorithm 2: Interactive segmentation

Input : Initial model $\mathcal{M}_0$, unlabeled image $\mathcal{I}$, number of iterations $\mathcal{N}$, feedback operation $\mathcal{R}$, conversion operation $\mathcal{C}$

Output: Updated model $\mathcal{M}_\mathcal{N}$

for $i \leftarrow 1$ **to** $\mathcal{N}$ **do**

    `/* generate segmentation map */`

    $\mathcal{S}_i \leftarrow \mathcal{M}_{i-1}(\mathcal{I})$;

    `/* get feedback from an expert */`

    $\mathcal{F}_i \leftarrow \mathcal{R}(\mathcal{S}_i, \mathcal{I})$;

    `/* convert to a new annotation */`

    $\mathcal{A}_i \leftarrow \mathcal{C}(\mathcal{F}_i)$;

    $\mathcal{M}_i \leftarrow$ fine-tuning $\mathcal{M}_{i-1}$ with $\mathcal{A}_i$;

end

return $\mathcal{S}_\mathcal{N}$

Leveraging Unlabeled Data

- 自我监督预训练
- 自监督
- 半监督

自我监督预训练

Comparison between self-supervised training methods that can directly or indirectly aid medical image segmentation.

Network	Surrogate task		
	Type	Description	Annotation
Encoder	Image-to-scalar	Predict if two longitudinal studies belong to the same patient	1(same)/0(different)
Encoder	Image-to-scalar	Predict the order of two slices random selected from the same CT scan	0(top)/1(bottom)
Encoder	Image-to-scalar	Predict the degree of rotation applied to a chest CT scan	$\frac{\theta}{90^\circ}$ $(\theta \in \{0, 90, 180, 270\}$
Siamese	Image-to-scalar	Predict the distance between two patches sampled from the same MR image	Float distance
Siamese	Image-to-scalar	Predict if two patches sampled from the same MR image are spatially near	1(near)/0(far)
Encoder-decoder	Image-to-image	Learn how to remove noise from MR image patches	Original patch before injec
Encoder-decoder	Image-to-image	Learn how to colorize gray-scale colonoscopy frames	Original frame before remo
Encoder-decoder	Image-to-image	Learn how to colorize gray-scale tele-med skin images	Original image before rem
Encoder-decoder	Image-to-image	Learn how to restore the image from various degradation transformations	Original image before degr
Encoder-decoder	Image-to-image	Learn how to weakly localize anatomical landmarks in MR images	Approximate landmark pos

自学习

Algorithm 3: Self learning

Input : Small labeled dataset $\mathcal{L}$, unlabeled dataset $\mathcal{U}$, iteration times $\mathcal{T}$, masks generation function $\mathcal{F}$

Output: Updated model $\mathcal{M}_{\mathcal{T}}$

$\mathcal{M}_0 \leftarrow$ training base model with $\mathcal{L}$;

for $i \leftarrow 1$ **to** $\mathcal{T}$ **do**

/* generate pseudo segmentation masks */

$\mathcal{S}_i \leftarrow \mathcal{F}(\mathcal{M}_{i-1}, \mathcal{U})$;

$\mathcal{D}_i \leftarrow \mathcal{L} \cup \{(\mathbf{x}, s) | \mathbf{x} \in \mathcal{U}, s \in \mathcal{S}_i\}$;

$\mathcal{M}_i \leftarrow$ fine-tuning $\mathcal{M}_{i-1}$ using $\mathcal{D}_i$;

end

return $\mathcal{M}_{\mathcal{T}}$

on	Initial annotations by	Pseudo masks generated by	Label noise handled by
2018a)	K-means	single segmentation model	N/A
017)	expert	single segmentation model + CRF	N/A
2018a)	expert	ensemble segmentation model	N/A
2018)	expert	ensemble segmentation model	a two-stream network
018)	expert	single segmentation model	a discriminator network

半监督学习

Algorithm 4: Semi-supervised learning

Input : Limited labeled dataset $\mathcal{L}$, unlabeled dataset $\mathcal{U}$, shared backbone $\mathcal{M}_c$, branch model and loss function for labeled data $\mathcal{M}_l, \ell_l$, branch model and loss function for unlabeled data $\mathcal{M}_u, \ell_u$

Output: Fine-tuned model $\mathcal{M}$

$\zeta_l \leftarrow \ell_l(\mathcal{M}_l(\mathcal{M}_c(\mathcal{L})))$;

$\zeta_u \leftarrow \ell_u(\mathcal{M}_u(\mathcal{M}_c(\mathcal{U})) + \ell_u(\mathcal{M}_u(\mathcal{M}_c(\mathcal{L}))$;

minimize$(\zeta_l + \zeta_u)$;

return $\mathcal{M}$

Publication	Unsupervised task
Bai et al. (2017)	Embedding consistency
Zhang et al. (2017b)	Image classification
Sedai et al. (2017)	Image reconstruction
Baur et al. (2017)	Manifold learning
Chartsias et al. (2018)	Image reconstruction
Huo et al. (2018a)	Image synthesis
Zhao et al. (2019)	Image registration
Li et al. (2019)	Transformation consistency