



HAL
open science

An unsupervised domain adaptation brain CT segmentation method across image modalities and diseases

Daqiang Dong, Guanghui Fu, Jianqiang Li, Yan Pei, Yueda Chen

► **To cite this version:**

Daqiang Dong, Guanghui Fu, Jianqiang Li, Yan Pei, Yueda Chen. An unsupervised domain adaptation brain CT segmentation method across image modalities and diseases. *Expert Systems with Applications*, 2022, 207, 10.1016/j.eswa.2022.118016 . hal-03760531

HAL Id: hal-03760531

<https://inria.hal.science/hal-03760531>

Submitted on 27 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Unsupervised Domain Adaptation Brain CT Segmentation Method Across Image Modalities and Diseases

Daqiang Dong^a, Guanghui Fu^{a,b,c,d,e,f,g}, Jianqiang Li^{a,*}, Yan Pei^h and Yueda Chenⁱ

^aBeijing University of Technology, Beijing, China

^bSorbonne Université, Paris, France

^cInstitut du Cerveau - Paris Brain Institute - ICM, Paris, France

^dInserm, Paris, France

^eCNRS, Paris, France

^fAP-HP, Hospital de la Pitié Salpêtrière, Paris, France

^gInria, Aramis project-team, Paris, France

^hComputer Science Division, University of Aizu, Aizuwakamatsu, Fukushima, Japan

ⁱTianjin Huanhu Hospital, Tianjin, China

ARTICLE INFO

Keywords:

Intraparenchymal hemorrhage
Semantic segmentation
Domain adaptation
Image synthesis
Deep learning

ABSTRACT

Computed tomography (CT) is the primary diagnostic tool for brain diseases. To determine the appropriate treatment plan, it is necessary to ascertain the patient's bleeding volume. Automatic segmentation algorithms for hemorrhagic lesions can significantly improve efficiency and avoid treatment delays. However, for deep supervised learning algorithms, a large amount of labeled training data is usually required, making them difficult to apply clinically. In this study, we propose an unsupervised domain adaptation method that is an unsupervised domain adaptation segmentation model that can be trained across modalities and diseases. We call it AMD-DAS for brain CT hemorrhage segmentation tasks. This circumvents the heavy data labeling task by converting the source domain data (MRI with glioma) to our task's required data (CT with Intraparenchymal hemorrhage (IPH)). Our model implements a two-stage domain adaptation process to achieve this objective. In the first stage, we train a pseudo-CT image synthesis network using the CycleGAN architecture through a matching mechanism and domain adaptation approach. In the second stage, we use the model trained in the first stage to synthesize the pseudo-CT images. We use the pseudo-CT with source domain labels and real CT images to train a domain-adaptation segmentation model. Our method exhibits a better performance than the basic one-stage domain adaptation segmentation method (+11.55 Dice score) and achieves an 86.93 Dice score in the IPH unsupervised segmentation task. Our model can be trained without using a ground-truth label, therefore increasing its application potential. Our implementation is publicly available at <https://github.com/GuanghuiFU/AMD-DAS-Brain-CT-Segmentation>

1. Introduction

Stroke is the second leading cause of death among adults worldwide, and one in four people will be affected throughout their lives (GBD 2016 Lifetime Risk of Stroke Collaborators, 2018). Intraparenchymal haemorrhage (IPH) accounts for 6.5% – 19.6% stroke cases (Feigin et al., 2003; O'donnell et al., 2010), but is associated with the highest rate of mortality (Gross et al., 2019). IPH is more likely to result in death or major disability than ischaemic stroke or subarachnoid hemorrhage and therefore constitutes an immediate medical emergency.


Computed tomography (CT) is typically used for the diagnosis, treatment planning, and prognosis monitoring of patients with IPH. Traditionally, radiologists visualize hemorrhage by artificially delineating CT scans and estimating its initial volume through the tada (ABC/2) formula (Kothari et al., 1996). This is the primary measurement for determining a follow-up treatment plan. Due to the urgency

of the onset of IPH, we hope to obtain the contour of the hemorrhage as quickly as possible. Therefore, accurate automated segmentation is essential to first aid.

Recently, deep-learning-based methods have achieved significant results in the domain of image segmentation. In particular, the models represented by U-net (Ronneberger et al., 2015) and its variants (Çiçek et al., 2016; Zhou et al., 2018; Xiao et al., 2018; Oktay et al., 2018; Ibtehaz and Rahman, 2020) have achieved excellent performance in medical imaging segmentation tasks. However, these methods are using supervised learning, which requires a large amount of pixel-level labeled data. Due to many reasons, e.g., privacy protection and expensive and professional labels, the medical data are difficult to obtain to train the learning model. Efficient transfer of available knowledge is a feasible approach to address these challenges. Specifically, domain adaptation methods can adapt to various data distributions in an unsupervised or semi-supervised manner and effectively use source domain data (Tzeng et al., 2017).

We consider the aforementioned unsupervised automated segmentation challenge as a domain adaptation problem. In this way, we can make use of the knowledge in

*Corresponding author

 dongdaqiang@emails.bjut.edu.cn (D. Dong);

guanghui.fu@icm-institute.org (G. Fu); lijianqiang@bjut.edu.cn (J. Li);

peiyan@u-aizu.ac.jp (Y. Pei); cyd681123@163.com (Y. Chen)

ORCID(s):

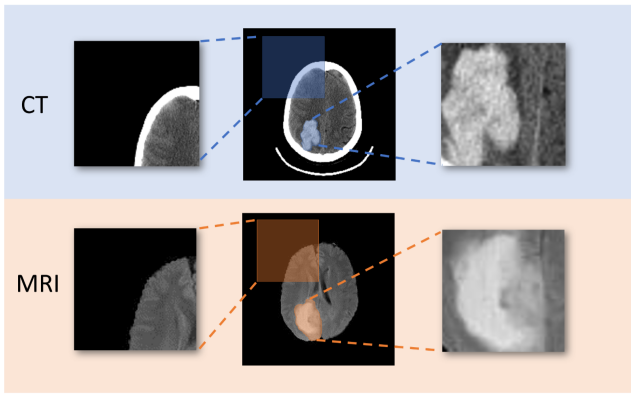


Figure 1: The difference between the source domain and target domain (MRI vs. CT): the manifestations of different diseases (glioma vs. cerebral hemorrhage) and the presence of the skull. We can see a noticeable difference from the enlarged part in this figure. The bright circular stripes in CT are CT head holders.

labeled data in similar domains. Specifically, we considered transferring knowledge from an open segmentation dataset of MRI with glioma through our proposed domain adaptation method to solve the problem of unsupervised IPH segmentation. In the defined problem, the images of the source domain (MRI with brain glioma) and the target domain (CT with IPH) are similar. IPH may mimic tumors on neuroimaging, both of which have irregular mass effects (Huisman and AGM, 2009), making the knowledge transferable intuitively. Domain adaptation methods in the medical domain focus on solving the variations between datasets. These include variations in imaging protocols (Kamnitsas et al., 2017), variations in various modalities (Chen et al., 2019), or diverse appearances of the same lesion in diverse datasets (Dong et al., 2020).

In our task, knowledge transfer spans not only modalities, but also type of diseases. So we need to build a model that crosses data modalities and disease types, and it is not ideal to use domain adaptation segmentation (DAS) alone. Because of the various types of diseases, not all source-domain images are conducive to our target tasks, e.g., lesions of different locations or distinct sizes. For this task, we propose a model called AMD-DAS. It is an unsupervised **Domain Adaptation Segmentation** model that can be trained **Across Modalities and Diseases**. To bridge the gap between data modalities and diseases, our model (AMD-DAS) comprises two stages to implement. In the first stage, we train a pseudo-CT image synthesis network to minimize the difference between the two modalities. In the second stage, we use labeled pseudo-CT images (obtained from the first-stage network) and unlabelled CT images to train the domain-adaptation segmentation network.

Our main contributions to this study are as follows.

- 1) We propose a domain adaptation image synthesis network that narrows the gap in the image appearance of modalities and retains the shape of lesions well.

- 2) We propose a domain adaptation segmentation network to successfully improve the cross-modality and cross-disease segmentation performance.
- 3) An efficient matching mechanism was proposed. The experiments show that it can reduce the negative effects of noisy data and improve performance in both the synthesis and segmentation stages.

Following this introduction section, we make a brief review in section 2 to introduce some works in segmentation fields. Our proposed AMD-DAS model is explained in detail in section 3. We present the two stages of implementation of this model. In section 4, we design experiments to evaluate our model and conduct a comparison with other models. The results and discussions are presented in section 5 and section 6, respectively. Finally, we conclude the whole work and give some open topics and remaining works for this study in the future in section 7.

2. Related Works

Our proposed method uses an image synthesis method to reduce the gap between the different modality appearances in the first stage. We use the domain adaptation method in the second stage to improve the segmentation performance of the target domain data. In this section, we analyze the semantic segmentation, image synthesis, and domain adaptation methods, which are related to our proposal.

2.1. Semantic Segmentation in Medical Domains

In recent years, many deep neural networks (DNN)-based methods have been attempted to solve automatic medical image segmentation problems, for example, (Pereira et al., 2016) and (Havaei et al., 2017) for brain tumour segmentation and (Chen et al., 2017) and (Karthik et al., 2019) for ischaemic lesion segmentation. Because of expensive data, supervised learning-based methods are not sufficiently flexible, and some unsupervised and semi-supervised learning methods have attracted many attentions.

Reference (Baur et al., 2020) considered the segmentation problem in brain MRI as an unsupervised anomaly detection problem, and they reconstructed an image through an auto-encoder to segment the lesion. (Bai et al., 2017) proposed a semi-supervised learning approach. This model is trained using both labeled and unlabelled data, and it improves the performance compared to the same amount of labeled data. Reference (Zhou et al., 2019) proposed a collaborative learning method to jointly improve the performance of disease grading and lesion segmentation by using semi-supervised learning with an attention mechanism. Reference (Xie et al., 2021) proposed an end-to-end scale-aware graph neural network (SAGNN) by reasoning the cross-scale relations among the support-query images for few-shot semantic segmentation.

The above methods contribute to the efficient usage of the target domain data. However, if the knowledge contained in the cross-domain data can be transferred, the application

potential of the model will be further enhanced. For example, although brain CT and brain MRI are cross-domains, they both represent brain structures. Relevant knowledge can be retained and transferred.

2.2. Style Transfer and Image Synthesis in Medical Domains

In the medical domain, many image synthesis methods are based on style transfers. Photographic style transfer methods transfer the style of a reference style photo onto another input photo. For instance, by appropriately choosing a reference style photo, one can make the input photo look like it was taken under different illumination, different times, different weather, or with another intention for artistic retouching (Luan et al., 2017).

Reference (Gatys et al., 2015) proposed a deep neural network that created artistic images of high perceptual quality. Subsequently, they proposed a neural algorithm of artistic style in (Gatys et al., 2016) that can separate and recombine the image content and transfer style. Reference (Zhu et al., 2017) proposed an unpaired image conversion network based on a generative adversarial network, called CycleGAN. This method has become the standard in this domain.

In the medical domain, (Lei et al., 2019) proposed dense cycle consistent generative adversarial networks using CycleGAN to generate CT images from MRI images. Subsequently, they utilized this method to synthesize an MRI and tried to assist in the diagnosis and treatment of the prostate in (Lei et al., 2020). (Jiang et al., 2018) introduced a method based on CycleGAN and adversarial domain adaptation to segment lung cancers. They presented tumor loss based on the additional supervision of masks to retain the tumor area of the synthetic images. For our task, as shown in Figure 1, transferring only the overall style between the source and target domain data is not enough. Because of the various types of diseases, we also need to consider retaining the semantic information of the source-domain data while transferring its modality characteristics.

2.3. Domain Adaptation

Domain adaptation methods are dedicated to solving the problem of domain shift between the source and the target domains. Reference (Ganin and Lempitsky, 2015) and (Ganin et al., 2016) proposed a domain-adversarial neural network (DANN) to narrow the distribution of feature representations. Reference (Hoffman et al., 2016) introduced the task of domain adaptation to semantic segmentation by applying adversarial learning in a fully convolutional manner to feature representations. In addition, to adapt feature representations, some methods have been used to adapt through segmentation masks, such as (Tsai et al., 2018). Reference (Hoffman et al., 2018) proposed a model that adapts between synthetic and real-world driving street-view domains using both generative image space alignment and latent representation space alignment. Reference (Xie et al., 2015) proposed a method for combining CNN with dictionary-based

models for scene recognition and visual domain adaptation. Reference (Chen et al., 2020) addressed the problems of adversarial training and feature distortion in unsupervised domain-adaptation semantic segmentation.

In the medical domain, reference (Zhang et al., 2020) proposed a novel collaborative unsupervised domain adaptation algorithm to conduct transfer ability-aware domain adaptation and cooperatively conquer label noise. Inspired by this work, we consider the transferability of the source domain images and integrated matching mechanism in our methods. Reference (Kamnitsas et al., 2017) presented a feature-adapted method for unsupervised domain adaptation in brain-lesion segmentation. They utilized knowledge from MRI of the same disease (traumatic brain injuries) but different modalities (gradient-echo to susceptibility-weighted images) for segmentation, which is limited in application scenarios.

3. Proposed Method: AMD-DAS

As shown in Figure 2, the proposed model comprises two stages in its framework. In the first stage, we train a pseudo-CT image synthesis network. We use a semantic retention mechanism to enable the model to achieve the best conversion for the target segmentation task. In the second stage, we use the pseudo-CT images generated by the first-stage model and its label and real CT images to train the domain-adaptation segmentation network. In both stages, we utilize a matching mechanism to selectively match the image pairs to effectively transfer the domain knowledge. Our implementation is publicly available at ¹

3.1. Stage 1: Training of Image Synthesis Network

As mentioned above, we try to produce MR images similar to CT images in overall style and form special lesion characteristics (e.g., representative IPH hematoma zone around the mass effect), while retaining the contour of the original lesion. To achieve the above objectives, we design an image synthesis network, as shown in Figure 3. The network makes several improvements using the CycleGAN architecture. The basic architecture of CycleGAN ensures cycle consistency, and it completes the transformation of image style while maintaining the contour. In addition, we introduce lesion consistency and feature space consistency to ensure that the generated image meets our expectations. We also introduce a matching mechanism to improve the synthesized images.

Specifically, we train the image synthesis network by optimizing the loss function in Eq. (1).

¹<https://github.com/GuanghuiFU/AMD-DAS-Brain-CT-Segmentation>.

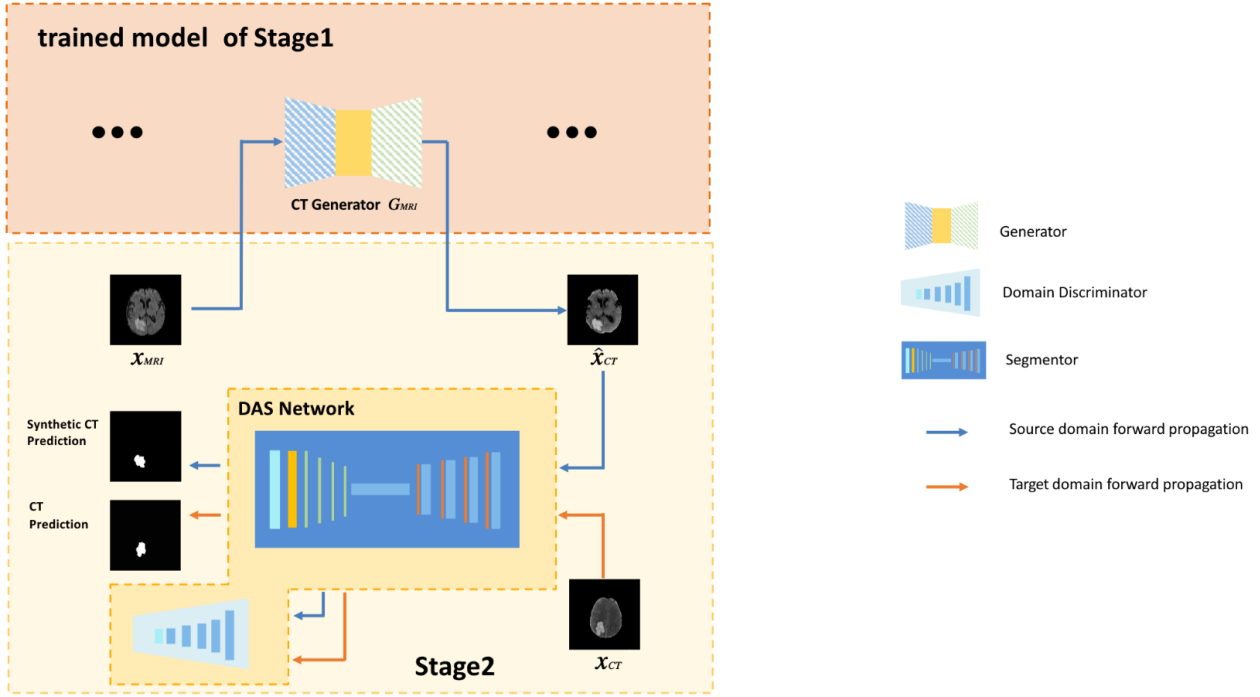


Figure 2: The overall architecture of our proposed method (for brevity, the skip connection in the segmented network is omitted)

$$\begin{aligned}
\mathcal{L} = & \mathcal{L}_{\text{GAN}}(G_{\text{MRI}}, D_{\text{CT}}) \\
& + \mathcal{L}_{\text{GAN}}(G_{\text{CT}}, D_{\text{MRI}}) \\
& + \lambda_{\text{cyc}} \mathcal{L}_{\text{cyc}}(G_{\text{MRI}}, G_{\text{CT}}) \\
& + \lambda_{\text{les}} \mathcal{L}_{\text{les}}(G_{\text{MRI}}, G_{\text{CT}}, S_{\text{ori}}, S_{\text{rec}}) \\
& + \lambda_{\text{fda}} \mathcal{L}_{\text{fda}}.
\end{aligned} \quad (1)$$

In Eq. (1), $\mathcal{L}_{\text{GAN}}(G_{\text{MRI}}, D_{\text{CT}})$ and $\mathcal{L}_{\text{GAN}}(G_{\text{CT}}, D_{\text{MRI}})$ are the adversarial losses proposed by CycleGAN (Zhu et al., 2017). The generator G_{MRI} attempts to confuse the discriminator D_{CT} by optimising the following adversarial loss, as shown in Eq. (2).

$$\begin{aligned}
\mathcal{L}_{\text{GAN}}(G_{\text{MRI}}, D_{\text{CT}}) \\
= & \mathbb{E}_{x_{\text{MRI}} \sim p_d(x_{\text{MRI}})} [\log(1 - D_{\text{CT}}(G_{\text{MRI}}(x_{\text{MRI}})))] \\
& + \mathbb{E}_{x_{\text{CT}} \sim p_d(x_{\text{CT}})} [\log(D_{\text{CT}}(x_{\text{CT}}))].
\end{aligned} \quad (2)$$

We update the discriminator by maximizing \mathcal{L}_{GAN} with a fixed G_{MRI} in Eq. (3).

$$D_{\text{CT}} = \underset{D_{\text{CT}}}{\text{argmax}} \mathcal{L}_{\text{GAN}}(G_{\text{MRI}}, D_{\text{CT}}). \quad (3)$$

The $\mathcal{L}_{\text{GAN}}(G_{\text{CT}}, D_{\text{MRI}})$ has nearly the same principle as $\mathcal{L}_{\text{GAN}}(G_{\text{MRI}}, D_{\text{CT}})$. We only write it once for brevity. The other three terms involve in the equation are described in detail in the following sections.

3.1.1. Cycle Consistency

This section briefly describes the basic principles of cycle consistency. To convert the basic style from MRI to CT, we first train the G_{MRI} module using an MRI x_{MRI} as input to generate a synthetic CT image. We train the G_{CT} module that uses the synthetic CT produced by the G_{MRI} as input to reconstruct the MRI. We minimize the cycle consistency loss to make the output image (reconstructed MRI) more similar to that of the original MRI. Referring to block a in Figure 3, the cycle consistency loss is defined as in Eq. (4).

$$\begin{aligned}
\mathcal{L}_{\text{cyc}}(G_{\text{MRI}}, G_{\text{CT}}) \\
= & \mathbb{E}_{x_{\text{MRI}} \sim p_d(x_{\text{MRI}})} [\|G_{\text{CT}}(G_{\text{MRI}}(x_{\text{MRI}})) - x_{\text{MRI}}\|_1] \\
& + \mathbb{E}_{x_{\text{CT}} \sim p_d(x_{\text{CT}})} [\|G_{\text{MRI}}(G_{\text{CT}}(x_{\text{CT}})) - x_{\text{CT}}\|_1].
\end{aligned} \quad (4)$$

Cycle consistency loss retains the basic semantic information between the original and reconstructed images by calculating their $L1$ distance.

3.1.2. Lesion Consistency

As shown in Figure 11, when we use CycleGAN for image synthesis, the output pseudo-CT can retain only the outline style. However, the key semantic information (lesion contour and texture) is lost, for example, through blurring and deformation of lesions. The cycle consistency loss is far from sufficiency to ensure that critical lesions are preserved.

We propose a lesion consistency loss for our problem, similar to the work in (Jiang et al., 2018). To transfer the lesion style from the glioma to IPH, we build two segmentation

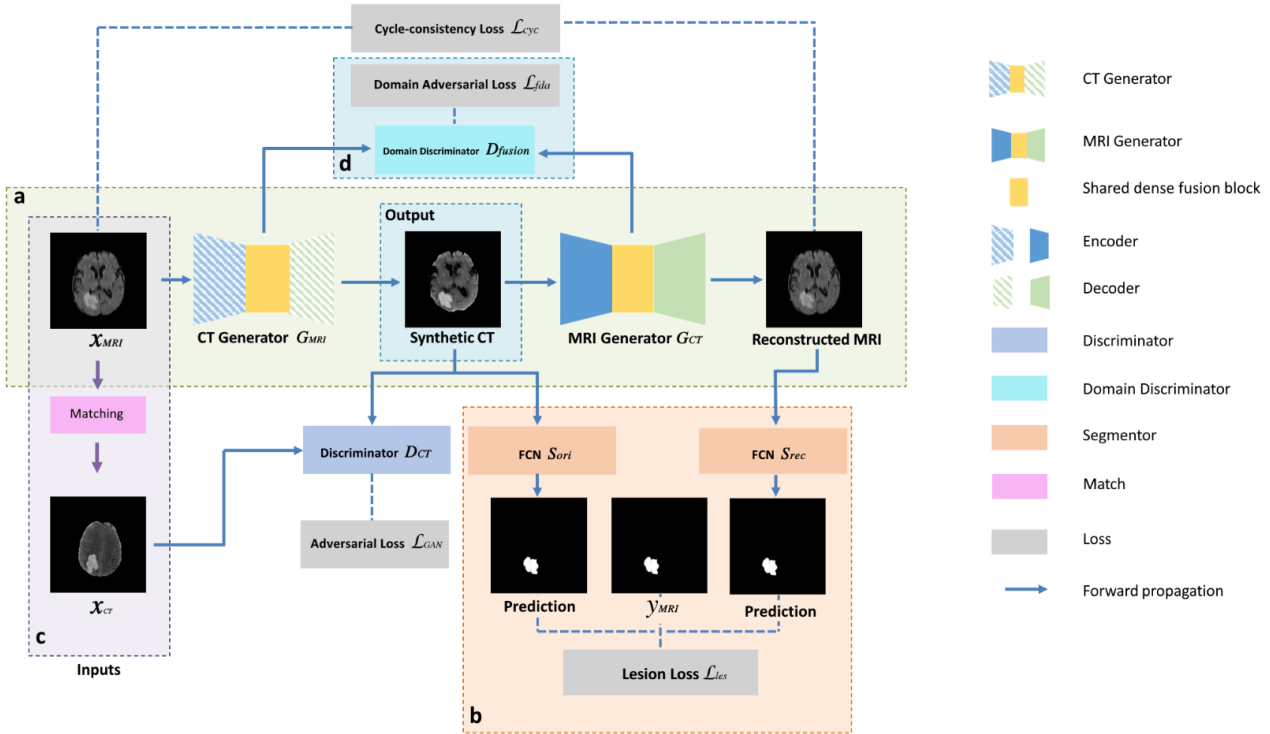


Figure 3: The architecture of the image synthesis network. Through the cycle consistency obtained from the reconstructed image, the lesion loss obtained by the auxiliary segmentation network, the image matching mechanism, and the feature fusion module with shared weights, we improve the performance of the synthetic image in the target task. The training process of the reverse loop is omitted. The four dashed blocks (a to d) correspond to the description in Section 3.1.

networks to predict the binary lesion mask of synthetic CT ($G_{MRI}(x_{MRI})$) and reconstructed MRI ($G_{CT}(G_{MRI}(x_{MRI}))$). Block **b** in Figure 3 depicts the principle of lesion consistency. We compare these two images with the true label y_{MRI} and to minimise the lesion consistency loss, as defined in Eq. (5):

$$\begin{aligned} \mathcal{L}_{les}(G_{MRI}, G_{CT}, S_{ori}, S_{rec}) \\ = \mathbb{E}_{x_{MRI} \sim p_d(x_{MRI})} [\ell(S_{ori}(G_{MRI}(x_{MRI})), y_{MRI}) \\ + \ell(S_{rec}(G_{CT}(G_{MRI}(x_{MRI}))), y_{MRI})]. \end{aligned} \quad (5)$$

ℓ in Eq. (5) is the cross-entropy loss function, and S_{ori} and S_{rec} are segmentation networks that predict the mask of the whole tumour. We optimize ℓ to update S_{ori} and S_{rec} .

3.1.3. Feature Space Domain Adaptation

Unlike CycleGAN, which uses two structurally identical but independent generators, our proposed generators consist of an encoder, a fusion module, and a decoder. The fusion module is a weight-sharing feature extractor. We consider the output of the G_{CT} 's encoder together with the fusion module as F_{CT} and the output of the G_{MRI} 's encoder together with the fusion module as F_{MRI} . Thus, the hidden layer representations of the images of these two modalities are mapped to the same subspace through the fusion module. Domain discriminator D_{fusion} is introduced in this module to

ensure the domain invariance of the extracted features, which is achieved by an adversarial training approach. Block **d** in Figure 3 depicts the above process. The adversarial and the discriminate losses of the feature space domain adaptation are represented as in Eq. (6):

$$\mathcal{L}_{fda}(x_{CT}) = - \sum_{h_f, w_f} \log(D_{fusion}(F_{CT})^{(h_f, w_f)}). \quad (6)$$

h_f, w_f is the size of the D_{fusion} 's output. This loss trains the fusion module and confuses the domain discriminator D_{fusion} by maximizing the probability of the target output being considered the source output.

F_{CT} and F_{MRI} are the outputs of the fusion module from the target and source domains, respectively, supervised by the discriminator D_{fusion} and are used to compute the cross-entropy loss \mathcal{L}_{d_f} as shown in Eq. (7), where $z = 0$ if the sample is drawn from the target domain and $z = 1$ for a sample from the source domain. In other words, we forward the F_{CT} and F_{MRI} using a cross-entropy loss \mathcal{L}_{d_f} for the target and source domains.

$$\begin{aligned} \mathcal{L}_{d_f} = - \sum_{h_f, w_f} (1 - z) \log(1 - D_{fusion}(F_{CT})^{(h_f, w_f)}) \\ + z \log(D_{fusion}(F_{MRI})^{(h_f, w_f)}). \end{aligned} \quad (7)$$

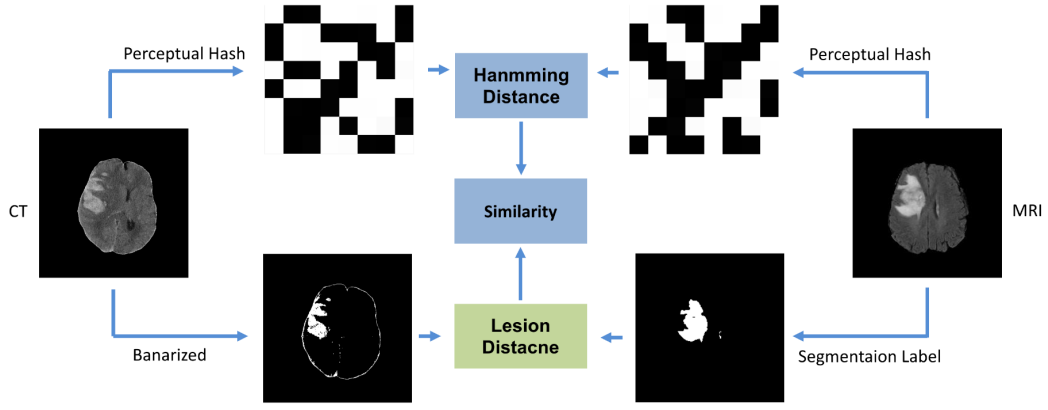


Figure 4: As shown in the figure for an MRI-CT pair, we calculate its similarity based on perceptual hash and the image after binarized.

3.1.4. Matching Mechanism

The original intention of the CycleGAN model transfers styles from the perspective of two sets of data. Refer to block **c** in Figure 3. Unlike CycleGAN, we use similar image pairs into the model during the training.

The motivations of this matching mechanism lie in the following two reasons.

1) Structural correlations between images from the target and source domains are crucial for image synthesis (Roy et al., 2014). The matching mechanism can select pairwise data with a high cross-domain correlation.

2) Selected highly correlated images benefit the generator and the discriminator. The generator learns to generate lesion areas instead of similar brain structures, and the matched images have highly correlated brain structures with the generated images, making it difficult to differentiate the discriminator. Consequently, it can help construct a more robust discriminator.

We call this process that inputs similar paired images during training the matching mechanism, and we propose a method to measure the similarity, as shown in Figure 4, which is described below.

The similarity of source and target domain images is calculated by two weighted distances: image Hamming distance and lesion distance. In the first part, we use the perceptual hash algorithm to calculate the image fingerprints of each image. The image fingerprint algorithm extracts image features and creates digital content as an identifier (Zou et al., 2015); (Lamon et al., 2001). See below for specific steps to calculate the similarity.

- First, we scale the original image (240×240) to $N \times N$ ($N = 32$) and note it as f , and perform a discrete cosine transform (DCT) on it to obtain the matrix F as shown in Eq. (8), Eq. (9) and Eq. (10). A denotes the DCT transform matrix.

$$F = AfA^T, \quad (8)$$

$$A(i, j) = c(i) \cos \left[\frac{(j + 0.5)\pi}{N} i \right], \quad (9)$$

$$c(i) = \begin{cases} \sqrt{\frac{1}{N}}, & i = 0, \\ \sqrt{\frac{2}{N}}, & i \neq 0. \end{cases} \quad (10)$$

- After that, we get the 8×8 sub-matrix B' in the top-left corner of F , as shown in Eq. (11) and Eq. (12)

$$F = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,32} \\ a_{2,1} & a_{2,2} & \dots & a_{2,32} \\ \vdots & \vdots & \dots & \vdots \\ a_{32,1} & a_{32,2} & \dots & a_{32,32} \end{bmatrix}, \quad (11)$$

$$B' = F[1, 8; 1, 8]. \quad (12)$$

- We calculated the mean value (\bar{a}) of the submatrix B' as shown in Eq. (13) and used it as a threshold to binarize B' to obtain the fingerprint \mathbf{b} of the image as shown in Eq. (14).

$$\bar{a} = \text{avg}(B'), \quad (13)$$

$$\mathbf{b} = \text{binary}(B', \text{threshold} = \bar{a}). \quad (14)$$

- Finally, we calculate the Hamming distance between the CT image's image fingerprint \mathbf{b}_{CT} and the MRI's image fingerprint \mathbf{b}_{MRI} as shown in Eq. (15)

$$d_{\text{ham}} = \mathbf{b}_{\text{MRI}} \oplus \mathbf{b}_{\text{CT}}. \quad (15)$$

\oplus denotes XOR calculation.

Then we calculate the lesion proportion of each image. For an MR image, we calculate the proportion of the tumour lesions as p_{tumour} . For a CT image, we binarize it through a threshold to roughly segment the lesion and calculate the proportion of IPH lesions as p_{iph} . The threshold is set to 120. We calculate the difference between p_{tumour} and p_{iph} as shown in Eq. (16).

$$d_{\text{les}} = |p_{\text{tumour}} - p_{\text{iph}}|. \quad (16)$$

We consider the weighted sum of the Hamming distance d_{ham} and lesion distance d_{les} as a measure of similarity, note it as d_{sim} , as shown in Eq. (17).

$$d_{\text{sim}} = \lambda_{\text{ham}} d_{\text{ham}} + d_{\text{les}}. \quad (17)$$

Note that the lesion distance d_{les} is the pixel difference between two images with the original resolution of 240×240 pixels, and in most cases, it is mainly contributed by the variations in lesion shapes. The Hamming distance $d_{\text{ham}} \in [0, 64]$ measures the difference in the level of the brain structures. The weighted measure d_{sim} balances them through λ_{ham} , a hyperparameter of our model chosen to be 100, with which we obtain our promising results.

In the training process of our model, we input matched image pairs, as shown in block c in Figure 3. For each MRI image, we sort the CT images according to its d_{sim} similarity with the MRI images and then selected the top n images ($n = 20$ in our training process) to form a one-to-many MRI and CT pair. This is to avoid the model learning a fixed mapping during backpropagation. An example of the MRI-CT pairs selected by our matching mechanism is shown in Figure 5.

3.2. Stage 2: Domain Adaptation Segmentation

In the medical domain, many methods use synthetic images regarding image synthesis as a means of data augmentation for weakly supervised learning (Jiang et al., 2018). However, no labels are available for our target domain images. Similar to the feature space domain adaptation in the synthesis network, we hope to project the features extracted by the segmentation network into the same subspace so that the hidden layer representations of the two domains are similar. We use the concept of an adversarial network to implement these goals. The overall definition is as in Eq. (18) and Eq. (19):

$$\hat{x}_{\text{CT}} = \tilde{G}_{\text{MRI}}(x_{\text{MRI}}), \quad (18)$$

$$\mathcal{L}_{\text{DAS}} = \mathcal{L}_{\text{seg}}(\hat{x}_{\text{CT}}) + \lambda_{\text{adv}} \mathcal{L}_{\text{adv}}(x_{\text{CT}}). \quad (19)$$

\tilde{G}_{MRI} in Eq. (18) is a trained generator G_{MRI} in stage 1.

First, we utilize a segmentation model fully convolutional network (FCN) (Long et al., 2015) S with a domain discriminator D_{seg} (the same architecture as the discriminator D_{fusion} in Section 3.1, shown in Figure 6). This

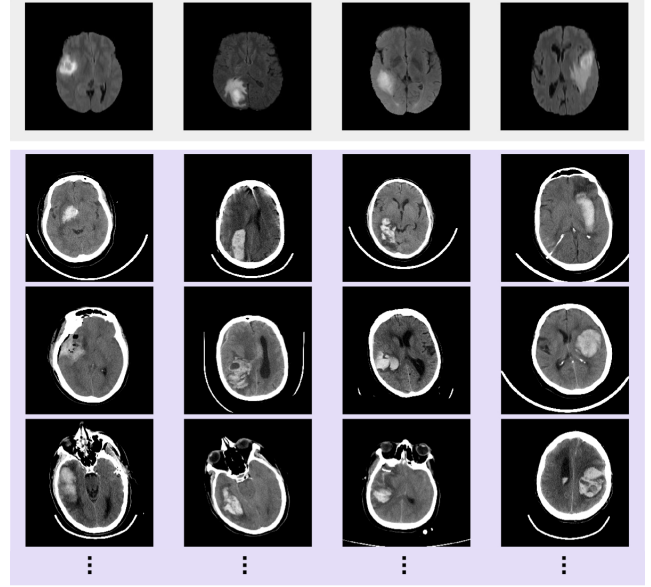


Figure 5: As shown in the figure, each column is an MRI and its corresponding similar CT images, which are matched by our method. We only show the top three CT images with the highest similarity in this figure. The bright circular stripes in CT are CT head holders.

module outputs a pixel-level segmentation map and a hidden layer representation. For the hidden layer output branch, regarded as a generator, we pass the hidden representation produced by this generator to the domain discriminator D_{seg} and optimize the GAN loss to fool D_{seg} . To optimize the performance of segmentation tasks in the source domain, we use a cross-entropy loss function as in Eq. (20).

$$\begin{aligned} \mathcal{L}_{\text{seg}}(\hat{x}_{\text{CT}}) = & - \sum_{h_o, w_o} y_{\text{MRI}}^{h_o, w_o} \log(\hat{P}(h_o, w_o)) \\ & + (1 - y_{\text{MRI}}^{h_o, w_o}) \log(1 - \hat{P}(h_o, w_o)). \end{aligned} \quad (20)$$

In Eq. (20), h_o and w_o represent the size of the predicted segmentation mask, which is equal to the size of the input image, $y_{\text{MRI}}^{h_o, w_o}$ is the pixel value of the true mask of the lesion, and \hat{P} is the segmentation mask output branch of the proposed FCN S . Through adversarial segmentation training, the feature map generated by the segmentation network fools the discriminator. The discriminator improves its discrimination ability through discriminative loss \mathcal{L}_D as in Eq. (21) and made their embedding similar.

$$\begin{aligned} \mathcal{L}_D = & - \sum_{h_f, w_f} (1 - z) \log(1 - D_{\text{seg}}(H_{\text{CT}})^{(h_f, w_f)}) \\ & + z \log(D_{\text{seg}}(\hat{H}_{\text{CT}})^{(h_f, w_f)}). \end{aligned} \quad (21)$$

Here, $z = 0$ if the sample is drawn from the target domain, and $z = 1$ for the sample from the source domain. H is the hidden-layer output branch of the proposed domain-adaptation segmentation network.

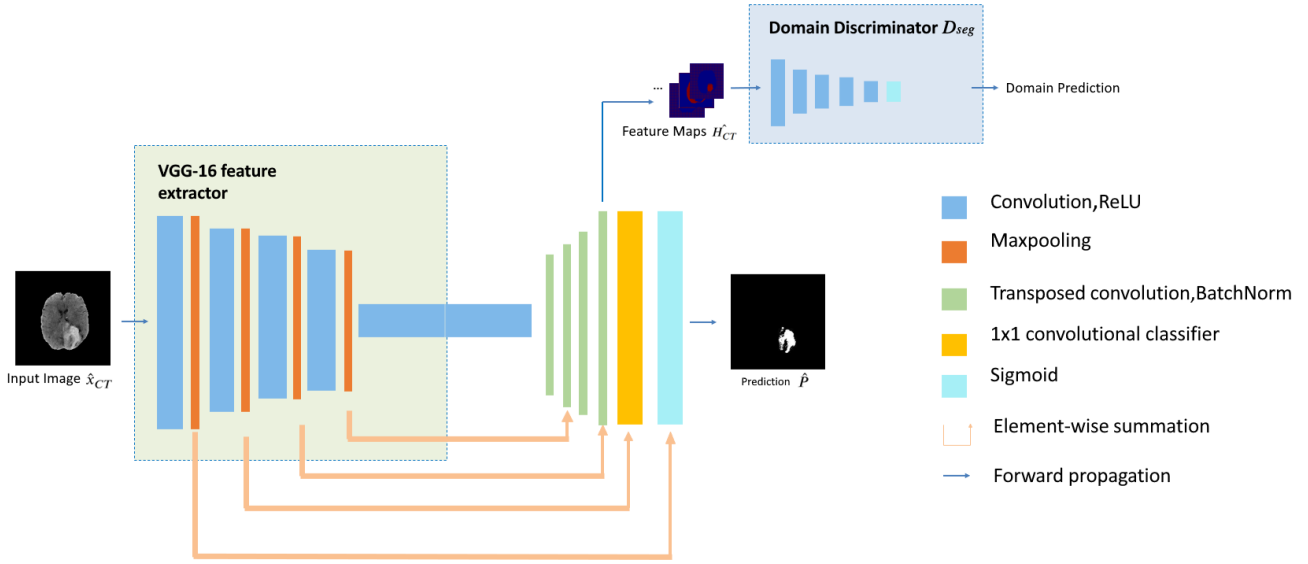


Figure 6: The domain adaptation segmentation part in our proposed model.

To narrow the distribution of the hidden layer representation to perform domain adaptation, we also use adversarial loss, as in Eq. (22).

$$\mathcal{L}_{\text{adv}}(x_{\text{CT}}) = - \sum_{h_f, w_f} \log(D_{\text{seg}}(H_{\text{CT}})^{(h_f, w_f)}). \quad (22)$$

We jointly optimize the adversarial and segmentation losses. Although we improve the segmentation performance of the model, the features extracted by embedding were domain-invariant. Through this mechanism shows in Eq. (23), the segmentation and cross-domain performances of the model are balanced.

$$\max_{D_{\text{seg}}} \min_S \mathcal{L}_{\text{DAS}}(S, D_{\text{seg}}). \quad (23)$$

We apply the matching mechanism to the domain-adaptation segmentation process, as described in Section 3.1.4. Not all source domain data are conducive to the domain-adaptation process for the target task. In this method, we select the corresponding MRI with the highest similarity for each CT image. Experiments have shown that the matching mechanism can help us use fewer source domain images to obtain better segmentation performance on target-domain tasks.

4. Evaluations

4.1. Dataset

In our study, the data are divided into two categories, i.e., source domain data and target domain data. The source domain data are MRIs of a brain tumor, and the target domain data are brain CT with IPH. Different slices from the same patient do not appear in the training or test set at the same time to avoid data leakage.

4.1.1. Source Domain Data

This dataset is used for the subtask of a decathlon medical segmentation challenge² (Simpson et al., 2019). The training set included 484 scans of $155 \times 4 \times 240 \times 240$ brain MR images. Each scan contained 155 images (a fixed resolution of 240×240) with four modalities: FLAIR, T1, T1c, and T2. In addition, three types of labels were provided to describe the various glioma subregions: enhancing tumor (ET), tumor core (TC), and whole tumor (WT). The four MR modalities and three types of labels can be seen in Figure 7. Among these, WT describes the full extent of the disease: tumor core and peritumoral edema (ED). It is typically depicted by a hyperintense signal on FLAIR. In CT images diagnosed as IPH, the hemorrhage often presents as a hyperintense signal (higher Hounsfield unit (HU) value). Therefore, we select FLAIR MRI and its whole tumor label as the source domain data pair. We select 387 sets of data as the training set and only took the 60th to 99th slices of each scan for a total of 15,480 slices. The lesion is easily visible in this range.

4.1.2. Target Domain Data

Target domain data are selected from RSNA's haemorrhage CT dataset³ (Flanders et al., 2020). Each CT series in this dataset has between 20 and 60 axial images that are 3 to 5 mm thick; the rows and columns are both 512. It contains a very large and heterogeneous collection of brain CT studies from multiple scanner manufacturers, institutions, and countries. A CT series starting from a reference line (Reid's baseline) is defined as continuous if the set contained all slices and discontinuous otherwise. We select 15,480 slices (discontinuous) in the training set without the pixel-level

²<http://medicaldecathlon.com/>

³<https://www.kaggle.com/c/rsna-intracranial-haemorrhage-detection/overview/>

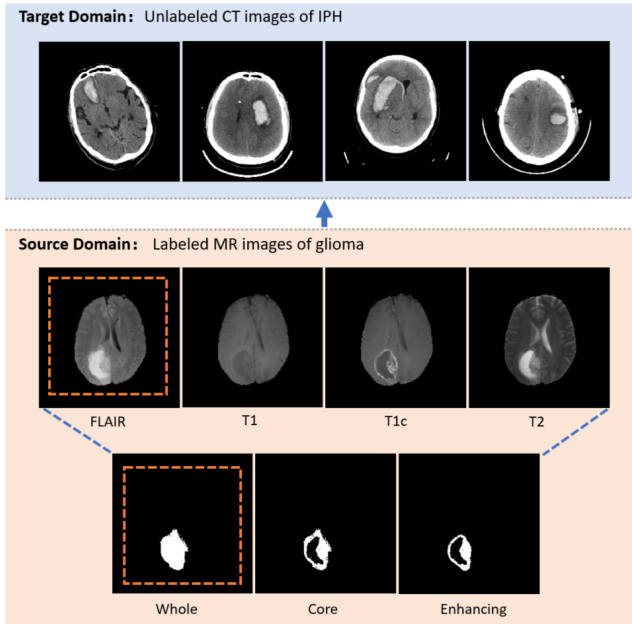


Figure 7: These two datasets span different image modalities (MRI and CT) and diseases (glioma and IPH), which is a tough challenge. In our task, we choose FLAIR modal MRI and a label covering the entire tumor as the source domain data pair (marked with the red dashed boxes). The bright circular stripes in CTs are CT head holders.

Table 1
Test dataset distribution

Disease	Number of slices
IPH	500
EDH	5
IVH	125
SAH	176
SDH	123
None	805

labels to ensure that many IPH clinical scenes were covered. An expert with 29 years of experience has labeled 1498 CT images for evaluation. In the unhealthy CT slices, we annotate 500 slices containing IPH at the pixel level (these slices may contain other diseases, we only annotate the pixel-level lesions of IPH and ignored other diseases). To evaluate the robustness of our method, we use 998 other slices as negative samples, including epidural (EDH), intraventricular (IVH), subarachnoid (SAH), and subdural (SDH), and healthy brain CT. The dataset distributions are listed in Table 1. Note that one CT scan may contain multiple diseases. All CT images are preprocessed using the bone removal method from (Muschelli et al., 2017) and re-sampled to a size of 240×240 pixels to match the source domain images.

4.2. Architecture and Training Details

4.2.1. Architecture

Generators: The generators consist of three parts: encoder, weight-shared fusion module, and decoder. The role

of the encoder is to downsample an image and obtain an initial feature map. The encoder first performs reflection padding, followed by a convolutional layer and instance normalization (Ulyanov et al., 2016). Subsequently, the following operations are repeated twice for downsampling: convolutional layers followed by instance normalization and the ReLU activation function. Inspired by the works of (Chen et al., 2018) and (Lei et al., 2019), the fusion module utilizes the dense connection proposed by (Huang et al., 2017) to perform the better feature extraction from previous and deeper hidden layers. The output part contains a deconvolution layer and \tanh activation layer to convert the feature representation into the prediction.

The decoder receives the encoder output and the fusion module's features as input and passes through the following module twice: a transposed convolution layer and a $ReLU$ activation layer. The output of the decoder contains a transposed convolution layer and a \tanh activation layer to convert the feature representation into a prediction.

Discriminators: D_{CT} and D_{MRI} The two discriminators D_{CT} and D_{MRI} in stage1 adopt the design of PatchGAN (Isola et al., 2017), which is a type of discriminator for generative adversarial networks that only penalizes structure at the scale of the local image patches. The discriminators start with a convolutional layer, followed by a leaky ReLU (0.2) (Maas et al., 2013). The following structure was repeated four times: a convolutional layer followed by a batch normalization process with a leaky ReLU (0.2). Finally, a convolutional layer is added to output the prediction.

Segmentation Network: For the segmentation network in both stages, we use a fully convolutional network and VGG-16 presented by (Simonyan and Zisserman, 2015) as the backbone for feature extraction. We use a series of fully convolutional layers and skipped connections (element-level addition) to aggregate abstract and concrete features.

The Domain Discriminator: Inspired by (Tsai et al., 2018), we utilize a domain discriminator used in the both stages (D_{fusion} and D_{seg}). It consists of fully convolutional layers to retain spatial information. The network consists of five convolution layers. For the last layer, we use sigmoid as the activation function, and for the other layers, we use leaky ReLU (0.2).

4.2.2. Training Details

We achieve adversarial and domain adaptation by alternately updating the network parameters in the following stages.

Training of the Synthesis Network (Stage 1): Each training pair contains MRI from the source domain and its corresponding labels, together with the matched CT images from the target domains. We set the batch size to two for training. The learning rates for optimizing the generators and discriminators are $2e-4$. All these are optimized using the Adam optimiser (Kingma and Ba, 2014) ($\beta_1=0.5$, $\beta_2=0.99$).

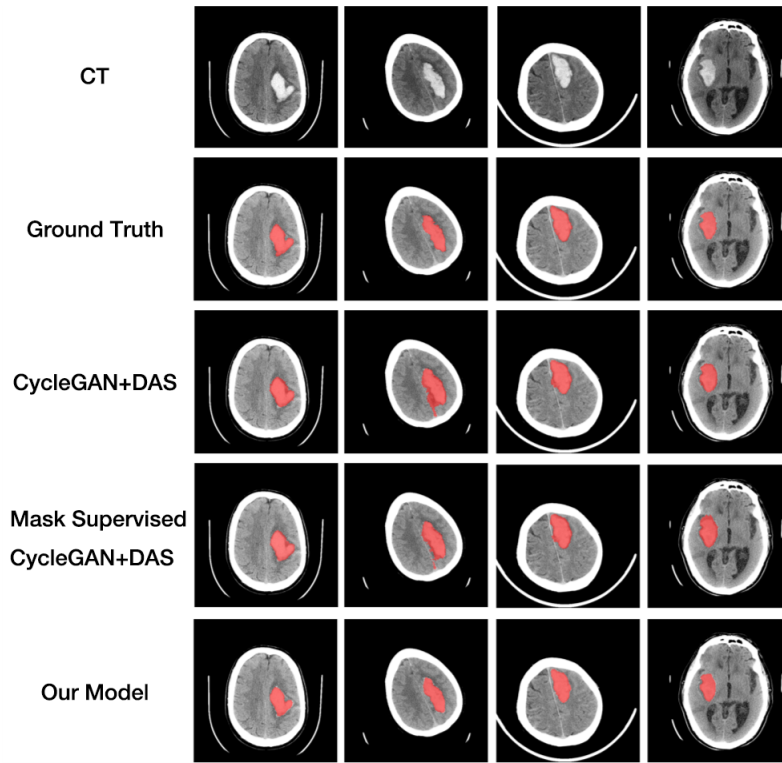


Figure 8: A comparison between the predictions (red region) of several models and the ground truth. We can see that our model surpasses other models in detail and achieves the best performance. The bright circular stripes in CTs are CT headholders.

The segmentation network was optimized with the SGD optimizer (momentum=0.7), and the learning rate was 0.1.

Training of the DAS Network (Stage 2): Each training sample contain a matched synthetic CT and its corresponding label, together with matched CT images. In each training batch, we first propagate the synthetic CT together with its label to our DAS network. Then, we forward-propagate the CT image to generate a feature map and optimize the domain discriminator. We set the batch size to 12 for training. The domain discriminator is optimized using the Adam optimizer ($\beta_1 = 0.9, \beta_2 = 0.99$), with a learning rate of 0.01, while the FCN is optimized using the SGD optimizer (momentum=0.7) with a learning rate of 0.1.

Loss functions: Our model optimizes a combination of five loss functions, as shown in Eq. (1). In our experiment, we set λ_{cyc} as 5, λ_{les} as 50 and λ_{fda} as 15. These hyperparameters are adjusted according to different tasks, data, and situations.

Evaluation metric: We use the Dice score as the evaluation metric. Using the definitions of true positive (TP), false positive (FP), and false-negative (FN), it is written as in Eq. (24):

$$DICE = \frac{2TP}{2TP + FP + FN}. \quad (24)$$

4.3. Ablation Experiments

4.3.1. Overview

We conduct ablation experiments from the following aspects:

1. We use various methods to synthesize pseudo-CT images from MRI and compare the results of training on pseudo-CT and testing on real CT.
2. We compare the performance of the one-stage and two-stage algorithms under the segmentation task.
3. The matching mechanism is evaluated in the aforementioned experiments. Furthermore, we construct five datasets to compare the performance of the network under different data volumes and distributions (whether matched) to evaluate the matching mechanism comprehensively.

4.3.2. Evaluation of Our Synthesis Network

Baseline: To evaluate the performance of our method, we define a baseline. We use the original MR images to train a segmentation model and evaluate the performance of the skull-stripped brain CT images.

CycleGAN: To evaluate the effect of using synthetic images, we use CycleGAN for image synthesis. We train the segmentation model using synthetic images (pseudo-CT images) and evaluate its performance.

Mask Supervised CycleGAN: To evaluate the effect after adding the lesion consistency loss, we add additional mask

Table 2
Evaluation of synthesis network

Method	Dice Score:%
Baseline	72.42
CycleGAN ((Zhu et al., 2017))	65.15
Mask supervised CycleGAN	67.67
Matched Mask supervised CycleGAN	76.28
The proposed synthesis network	81.41

supervision based on the CycleGAN architecture. We use this method to synthesize images and train the segmentation model. This method reflects the basic performance of a series of mask-based supervision methods, such as those proposed by (Jiang et al., 2018).

Matched Mask Supervised CycleGAN: To evaluate the effect of the proposed matching mechanism, we tested the mask-supervised CycleGAN after integration with the matching mechanism.

Our Synthesis Network: The overall architecture of our synthesis network is explained in Section 3.1. Based on the CycleGAN architecture and additional mask supervision, our method also adds the mapping of the same subspace at the feature level using a fusion module. By comparing our synthesis network and matched mask-supervised CycleGAN, the effect of the fusion module is verified.

The experiment results are listed in Table 2. The synthesized images of the different networks are shown in Figure 11.

4.3.3. Evaluation of Two-stage Methods

To measure the performance of stages 2 and 1 after integration and further verify the advantages of our synthesis method, we conducted the following experiments. All experiments performed data augmentation operations on the source domain data: random LR and AP flip (horizontal and vertical flip) with a probability of 50%, increasing or decreasing the brightness by 0.2.

DAS Network: First, with all source domain data used, after domain adaptation training, we obtained the segmentation model to predict the IPH region on CT images (with the skull).

Synthesis Network + DAS Network: We evaluated the performance of various synthetic networks combined with Stage 2. By comparing with the Dice score of stage 1 (shown in Table 2), the robustness of the domain adaptation segmentation network and superiority of the two-stage method were demonstrated.

The experimental results are listed in Table 3. Figure 8 shows the segmentation results of the above experiments.

Table 3
Evaluation of the whole models

Model	Stage	Dice Score:%
DAS	One	74.22
CycleGAN+DAS		72.31
Mask CycleGAN+DAS	Two	78.60
Matched Mask CycleGAN+DAS		80.14
The proposed model (AMD-DAS)		85.77

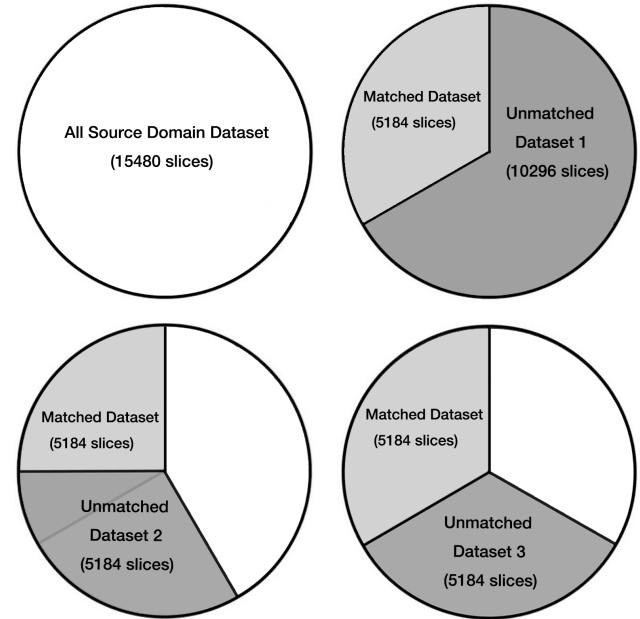


Figure 9: The figure shows the five datasets used to evaluate the matching mechanism (all source domain dataset, one matched dataset and three unmatched datasets).

4.3.4. Evaluation of the Matching Mechanism

The matching mechanism was applied in both stages. We can verify its effect by measuring the segmentation effect, which can be evaluated in Sections 4.3.2 and 4.3.3. To further verify its effect, we constructed the following datasets. Finally, we selected 5,184 images out of 15,480 source domain images (MRI) by the matching mechanism and called it the matched dataset. As listed in Table 4, we built three unmatched datasets. The relationship between these datasets is shown in Figure 9. Unmatched dataset 1 includes all the images in the source domain data except for the matched dataset, which contains 10,296 slices. Unmatched dataset 2 was randomly selected from all source domain data (it may contain images from the matched dataset). Unmatched dataset 3 was randomly selected from unmatched dataset 1 (which cannot contain the images in the matched dataset). The above two datasets maintained the same amount of data (5,184 slices) as the matched dataset.

We evaluated the segmentation performance of various source domain data using the same domain adaption segmentation model and the target domain data. The result can be seen in Table 4.

Table 4
Evaluation of the matching mechanism under different datasets

Dataset	Size	Dice score:%
All source domain data	15,480	85.77
Unmatched dataset 1	10,296	83.55
Unmatched dataset 2	5,184	83.24
Unmatched dataset 3	5,184	82.12
Matched dataset	5,184	86.93

Table 5
Comparison with Supervised Learning

Model	Pretrained	Dice score:%
FCN	None	93.82
	Instance: MRI	94.50
	Instance: Synthetic CT	95.01
	Parameters: AMD-DAS	95.10
U-Net	None	95.49
	Instance: MRI	95.73
	Instance: Synthetic CT	95.96

4.3.5. Performance as Transfer Learning

To explore the effect of the proposed method in supervised learning scenarios, we performed experiments to verify the two methods in Stage 1 (image synthesis) and Stage 2 (adaptation segmentation) as transfer learning strategies for supervised learning segmentation tasks. We constructed a 5-fold cross-validation by using 1498 CT images (500 labeled disease slices and 998 control slices) as the experimental dataset. Transfer learning methods can be categorized into four groups: instance-based, feature-based, parameter-based, and relational-based approaches (Pan and Yang, 2009; Zhuang et al., 2020), and we tried instance-based and parameter-based depending on our task and experiment setting. The instance-based transfer learning is to assume that certain parts of the data in the source domain can be reused for learning in the target domain by reweighting. We first trained FCN and U-Net as baseline models and experimented with instance-based transfer learning using MRI (Instance: MRI) or synthetic CT obtained from our Stage 1 method (Instance: Synthetic CT) as pre-training data. The parameter-based transfer learning assumes that the source tasks and the target tasks share some parameters or prior distributions of the hyperparameters of the models. We treat the overall trained unsupervised model as pre-training and fine-tune it with target CT to perform parameter-based transfer learning (parameter: AMD-DAS). Since the segmentation module of our model is based on FCN, we do this experiment on the FCN model. The experiment results can be seen in Table 5.

5. Results

Through the experimental results in Table 2, we found that for our specific tasks (adaptation for different types of diseases and different image modalities), the use of the

CycleGAN method resulted in a lower performance than the baseline (-7.27 Dice score). The segmentation task is highly dependent on the image context structure. The loss of lesion detail caused by this method severely weakens performance. Even with the addition of segmentation label supervision as mask-supervised CycleGAN, the segmentation score was slightly improved compared with CycleGAN (+2.52 Dice score), which is still not as good as the baseline (-4.75 Dice score). As shown in Figure 11, for the lesion part, the image synthesized by mask-supervised CycleGAN has a clear edge but is blurred inside. However, internal occupancy is crucial for the segmentation task. This is the reason for the poor performance. Subsequent experiments demonstrated that the introduction of the matching mechanism and domain adaptation in the feature space significantly improved the performance of the segmentation task and exceeded the baseline (+8.99 Dice score).

As shown in Table 3, applying the DAS network alone effectively improves the performance of the target task (beyond the baseline, +1.8 Dice score). A comparison with the corresponding single-stage method in Table 2 demonstrates the superiority of our two-stage method (shown in Figure 10). Figure 8 shows some of the segmentation results used for comparison.

Furthermore, the results in Tables 2, 3, and 4 proved that our matching mechanism is effective in both the synthesis and the DAS network. As shown in Table 4, generally the more source domain data we used, the better the performance: we used all source domain data to achieve better performance than using two-thirds (unmatched dataset 1) or one-third (unmatched dataset 2 or unmatched dataset 3) of the dataset. However, if we use the dataset selected by the matching algorithm, the best performance can be achieved. Under the same amount of data (unmatched datasets 2 and 3), the Dice score of 3.69 to 4.81 was increased. When the amount of data was doubled (unmatched dataset 1), the Dice score increased by 3.38; even in the case of three times the amount of data (all source domain datasets), the Dice score increased by 1.16.

The experimental results in Table 5 show that both the methods from stages 1 and 2 have benefits for supervised learning models by utilizing them as transfer learning. From an instance-based transfer learning perspective, using synthetic CT improves the Dice score by 0.51 for the FCN model and 0.23 for the UNet model compare with using MRI. After transfer learning, our method can improve the 1.19 dice score compared to the baseline FCN model and the 0.47 dice score compared to the UNet baseline. From the perspective of parameter-based transfer learning, utilizing our method as a pre-trained model can improve the Dice score by 1.28.

Efficiency is one of the most important indicators for making medical models applicable to clinics. Each scan in the RSNA dataset contains 60 CT slices. We run our experiments with NVIDIA Geforce GTX 1660Ti, and it cost 0.06 seconds on average for our model to obtain the segmentation result for each scan.

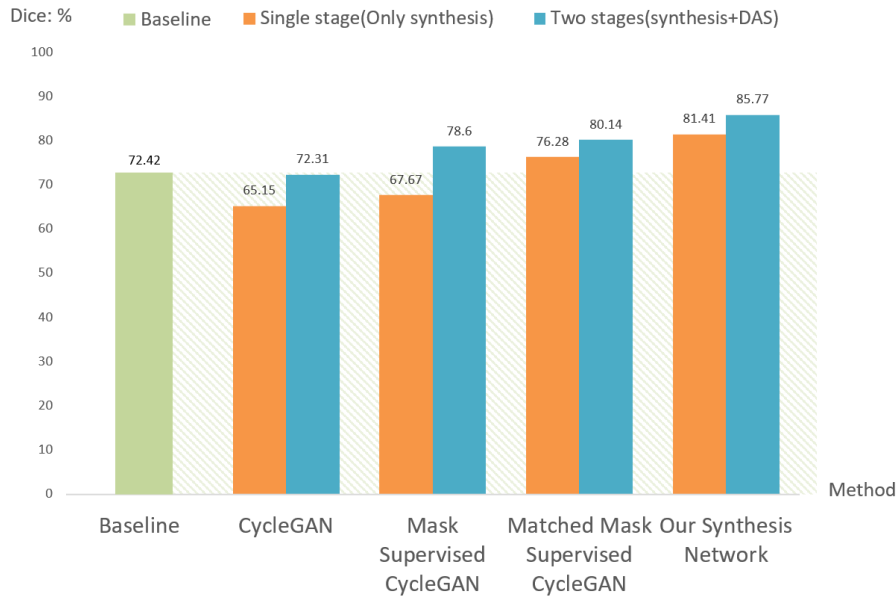


Figure 10: Dice scores of different methods.

6. Discussions and Analyses

6.1. Limitations on Different Subtypes of Intracerebral Hemorrhage

For patients with intracerebral hemorrhage (ICH), various hemorrhages often occur together. Surgical management needs to be determined in some clinical scenarios. Hematoma volume and location are the two main predictors of outcomes related to the hematoma itself (de Oliveira Manoel, 2020). The leading cause of nontraumatic SAH is the rupture of an intracranial aneurysm, which is not within the scope of our research. The publicly available MRI data with tumors have lesion shapes nearly similar to those of CT scans with IPH but are significantly different from those of subdural hemorrhage (SDH) and epidural hemorrhage (EDH). Therefore, we did not consider these two situations in the present study. Consequently, we have observed false positives (with a reduction of Dice scores) provoked by these lesions. From the perspective of clinical applicability and existing data, our research used data across modalities and disease types to design unsupervised IPH segmentation algorithms, which is a precedent in this field. In future studies, we will explore the application of this paradigm to different diseases based on clinical needs.

6.2. Effect of Matching Mechanism

To demonstrate the effectiveness of the matching mechanisms, we conducted a two-stage analysis. The first stage involved inputting paired images that were similar in terms of the location and structure of the hemorrhage. We found that this was effective in downstream segmentation, with the results showing a high Dice coefficient results. In Stage 2, we used a matching mechanism to selectively use the source domain data. The selected data included only the

higher-quality data in our dataset, which resulted in us using only one-third of the source data. We obtained a better performance by focusing on good-quality data rather than quantity.

Normally, training with more data will leads to better performance for a deep learning model. However, in this research, training with fewer data points selected by our matching mechanism leads to better performance. This is because some of the originally sourced data are riddled with noise and thus impact the performance. Note that complexity is still the main concern of image matching algorithms (Ma et al., 2021). Thus, we focused on a practical and efficient algorithm, instead of more complex algorithms, such as deep learning models, to design our matching mechanism. The primary limitation of the matching mechanism is that it is designed based on our dataset, which may lack generalization and affect its performance when measured against an external dataset. However, our algorithm has demonstrated its effectiveness based on various experiments. Such limitations can be eliminated by introducing normalization to be more easily adapted for other applications, which we will explore in our future work.

6.3. Unreal Generated Images in Experiments

In our experiments, basic data augmentation techniques are used, as in other related studies (Wang et al., 2017), (Chen et al., 2017). Most data augmentation methods may yield unrealistic images that violate anatomical structures. E.g., flipping vertically would show an upside-down brain scan; and our data synthesis methods also don't always produce perfect results.

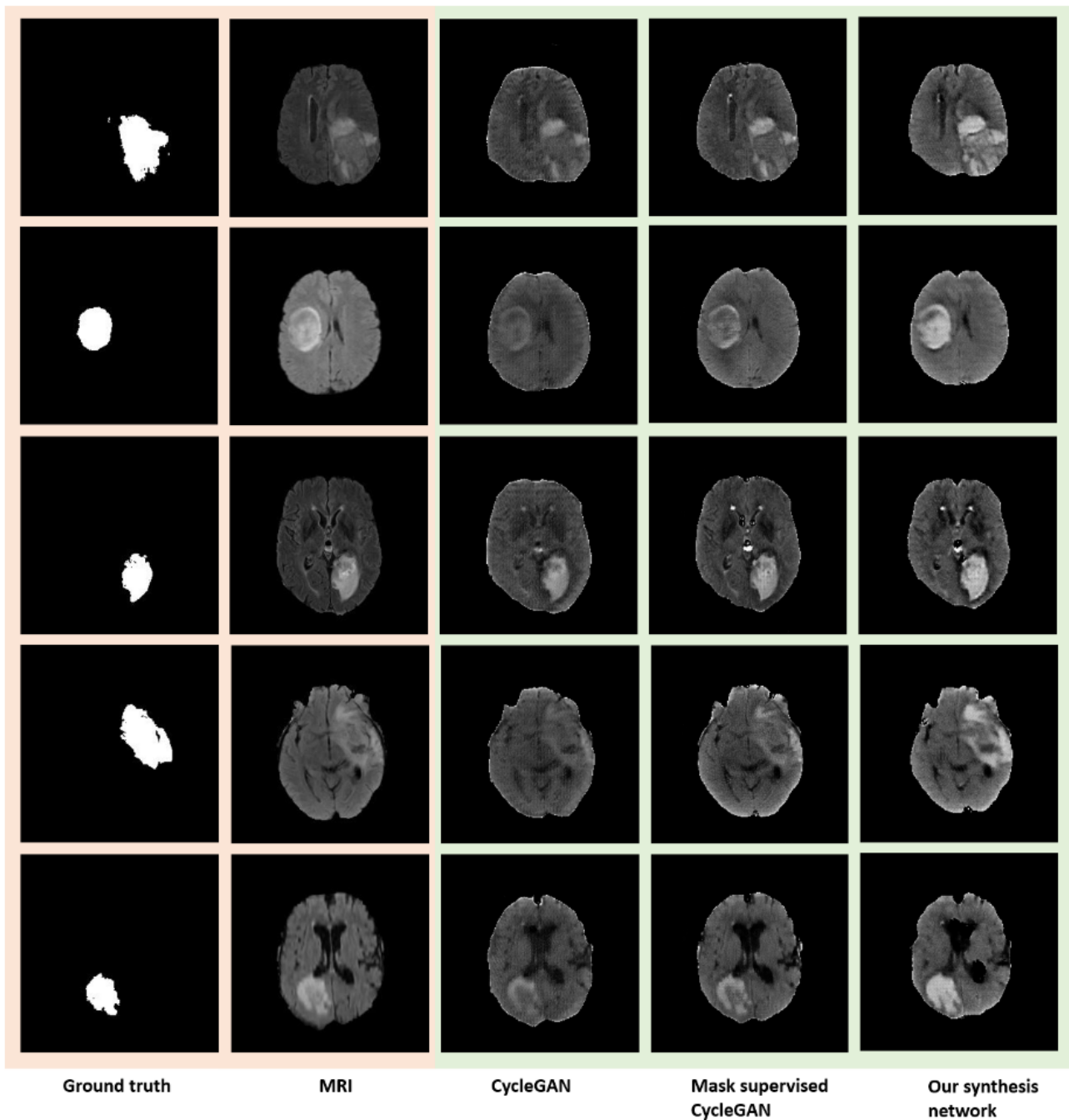


Figure 11: Example of synthetic images generated by various methods.

6.4. Consider the Correlation Between Slices

Model architectures that consider slice correlations can be broadly classified into two categories: 2D architectures with sequence-level encoder-decoder and 3D architectures at voxels-level. It is intuitively better to consider slice correlations, but we have to consider their high computational resource consumption. For instance, the image size has to be reduced to meet the resource requirements as (Korolev et al., 2017) and (Han et al., 2019) did when using 3D model. Therefore, 2D models that do not consider slice correlations are still a dominant research direction today. Our

view is that considering slice correlation is a gain for the model, and models without considering relations can also be used, but with care to avoid data leakage or underlying bias (Wallis and Buvat, 2022). Constructing a model that can consider slice correlation is one of our future works. The basic idea is to describe the characteristics of the variation between continuing slices. In multimodal medical imaging, the anatomical information contained is common despite the differences between contrast or resolution. Therefore, we believe that optical flow is a good way to characterize the anatomical information of the full sequence data. We will

try to consider optical flow for domain adaptation based on the proposed method in our future work.

6.5. One-stage versus the Two-stage Model

Our algorithm is divided into two stages and has better performance than the single-stage method based on the experimental results. Some studies have pointed out that the one-stage method may be beneficial, as illustrated in studies by (Ma et al., 2019) and (AlBadawi and Fraz, 2018). The authors of these studies claimed that in the two-stage paradigm, the errors produced during the first stage accumulate in the second stage, while the one-stage method avoids this disadvantage. However, to date, the two-stage method has been the mainstream method in this field with stable performance, for example, (Chartsias et al., 2017), (Zou et al., 2020) and (Chen et al., 2019). As the two-stage method can obtain the intermediate results more intuitively in the first stage, so it is easier to train. According to our experiments, the synthesized images in the first stage can also be used as instance-based transfer learning to improve the performance of the supervised algorithm, which also proves the value of this module. In the CrossMoDa 2021 challenge, (Dorent et al., 2022), image-to-image translation methods have the highest proportion and are currently the top method of this type. This proves that the type of model adopted in this study is currently popular and stable. Cross-domain medical data such as brain CT and brain MRI represent the same brain structure. If the method of disentangled representation can be used, it is believed that unnecessary generation can be reduced, and the transformation can be performed directly, making it more controllable. We will attempt to simplify our model based on the idea of disentangled representation and maintain performance improvement in future work.

6.6. In Supervised Learning Scenario

- Supervised Model for Our Task: Since tumors and intraparenchymal hemorrhage are both manifest as mass effects on CT slices, we have a public dataset (Simpson et al., 2019) of tumor segmentation annotations on MRI. Therefore, for our task, the supervised method is to train on MRI with tumor labels and directly verify the segmentation task of the IPH CT data, represented as a baseline algorithm. Compared with the baseline domain adaption segmentation algorithm (DAS), our model has improved performance (+11.55 Dice score).
- Supervised model from a traditional perspective: As described in Section 4.3.5, we trained several supervised learning models in five-fold cross-validation using the test set. Undoubtedly supervised learning achieves better performance than ours; however, our model has the critical advantage that it can be trained without data annotation. We found that the two-stage methods can be used as instance-based and parameter-based transfer learning modules to enhance the performance of supervised learning models. This proves that our two-stage model has a wide range of applications.

7. Conclusion

In this study, we proposed a two-stage brain CT segmentation model based on a domain-adaptation method. Our proposed model was trained using the source domain data of different disease types and image modalities to segment the target domain images. The synthesized images generated by our synthesis network were more conducive to the subsequent segmentation task than the other methods. Domain adaptation was performed for the segmentation task on the hidden layer in the feature space. Based on the characteristics of the task, we defined a method for measuring the similarity between the source and target images and proposed an image-matching mechanism. Experiments have shown that using only one-third of the matched data can achieve better performance than using all the data. Our model did not require true labels in the target domain, but efficiently transferred source-domain knowledge, and had good application potential.

Acknowledgments

The study is supported by the National Key R&D Program of China with project no. 2020YFB2104402 and Beijing Chaoyang Collaborative Innovation project with no. CYXC2010. Guanghui Fu is supported by the Chinese Government Scholarship provided by China Scholarship Council (CSC). The authors are grateful to Janan Arslan and Yue Ma for their comments and modification to the manuscript.

References

- AlBadawi, S., Fraz, M., 2018. Arterioles and venules classification in retinal images using fully convolutional deep neural network, in: International Conference Image Analysis and Recognition, Springer. pp. 659–668. URL: https://link.springer.com/chapter/10.1007/978-3-319-93000-8_75.
- Bai, W., Oktay, O., Sinclair, M., Suzuki, H., Rajchl, M., Tarroni, G., Glocker, B., King, A., Matthews, P.M., Rueckert, D., 2017. Semi-supervised learning for network-based cardiac mr image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 253–260. URL: https://link.springer.com/chapter/10.1007/978-3-319-66185-8_29, doi:https://doi.org/10.1007/978-3-319-66185-8_29.
- Baur, C., Denner, S., Wiestler, B., Albarqouni, S., Navab, N., 2020. Autoencoders for unsupervised anomaly segmentation in brain mr images: A comparative study. arXiv preprint arXiv:2004.03271 URL: <https://arxiv.org/abs/2004.03271>.
- Chartsias, A., Joyce, T., Dharmakumar, R., Tsaftaris, S.A., 2017. Adversarial image synthesis for unpaired multi-modal cardiac data, in: International workshop on simulation and synthesis in medical imaging, Springer. pp. 3–13. URL: https://www.pure.ed.ac.uk/ws/portalfiles/portal/44085363/SASHIMI_2017_final.pdf.
- Chen, C., Dou, Q., Chen, H., Qin, J., Heng, P.A., 2019. Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation, in: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 865–872. URL: <https://arxiv.org/abs/1901.08211>.
- Chen, L., Bentley, P., Rueckert, D., 2017. Fully automatic acute ischemic lesion segmentation in dwi using convolutional neural networks. NeuroImage: Clinical 15, 633–643. URL: <https://www.sciencedirect.com/science/article/pii/S221315821730147X>, doi:<https://doi.org/10.1016/j.nicl.2017.06.016>.

- Chen, L., Wu, Y., DSouza, A.M., Abidin, A.Z., Wismüller, A., Xu, C., 2018. Mri tumor segmentation with densely connected 3d cnn, in: *Medical Imaging 2018: Image Processing*, International Society for Optics and Photonics. p. 105741F. URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10574/105741F/MRI-tumor-segmentation-with-densely-connected-3D-CNN/10.1117/12.2293394>. short, doi:<https://doi.org/10.1117/12.2293394>.
- Chen, T., Zhang, J., Xie, G.S., Yao, Y., Huang, X., Tang, Z., 2020. Classification constrained discriminator for domain adaptive semantic segmentation, in: *2020 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE. pp. 1–6. URL: <https://ieeexplore.ieee.org/document/9102965>.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3d u-net: learning dense volumetric segmentation from sparse annotation, in: *International conference on medical image computing and computer-assisted intervention*, Springer. pp. 424–432. URL: https://link.springer.com/chapter/10.1007/978-3-319-46723-8_49, doi:https://doi.org/10.1007/978-3-319-46723-8_49.
- Dong, J., Cong, Y., Sun, G., Zhong, B., Xu, X., 2020. What can be transferred: Unsupervised domain adaptation for endoscopic lesions segmentation, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4023–4032. URL: <https://arxiv.org/abs/2004.11500>.
- Dorent, R., Kujawa, A., Ivory, M., Bakas, S., Rieke, N., Joutard, S., Glocker, B., Cardoso, J., Modat, M., Batmanghelich, K., et al., 2022. Crossmoda 2021 challenge: Benchmark of cross-modality domain adaptation techniques for vestibular schwannoma and cochlea segmentation. arXiv preprint arXiv:2201.02831 URL: <https://arxiv.org/abs/2201.02831>.
- Feigin, V.L., Lawes, C.M., Bennett, D.A., Anderson, C.S., 2003. Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *The lancet neurology* 2, 43–53. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1474442203002667>, doi:[https://doi.org/10.1016/S1474-4422\(03\)00266-7](https://doi.org/10.1016/S1474-4422(03)00266-7).
- Flanders, A.E., Prevedello, L.M., Shih, G., Halabi, S.S., Kalpathy-Cramer, J., Ball, R., Mongan, J.T., Stein, A., Kitamura, F.C., Lungren, M.P., et al., 2020. Construction of a machine learning dataset through collaboration: the rsna 2019 brain ct hemorrhage challenge. *Radiology: Artificial Intelligence* 2, e190211. URL: <https://pubs.rsna.org/doi/10.1148/ryai.2020190211>.
- Ganin, Y., Lempitsky, V., 2015. Unsupervised domain adaptation by backpropagation, in: *International conference on machine learning*, pp. 1180–1189. URL: <http://proceedings.mlr.press/v37/ganin15.html>.
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., Lempitsky, V., 2016. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research* 17, 1–35. URL: <http://jmlr.org/papers/v17/15-239.html>.
- Gatys, L.A., Ecker, A.S., Bethge, M., 2015. A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576 URL: <https://arxiv.org/abs/1508.06576>.
- Gatys, L.A., Ecker, A.S., Bethge, M., 2016. Image style transfer using convolutional neural networks, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2414–2423. URL: https://openaccess.thecvf.com/content_cvpr_2016/html/Gatys_Image_Style_Transfer_CVPR_2016_paper.html.
- GBD 2016 Lifetime Risk of Stroke Collaborators, 2018. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. *New England Journal of Medicine* 379, 2429–2437. URL: <https://www.nejm.org/doi/full/10.1056/NEJMoa1804492>.
- Gross, B.A., Jankowitz, B.T., Friedlander, R.M., 2019. Cerebral intraparenchymal hemorrhage: a review. *Jama* 321, 1295–1303. URL: <https://jamanetwork.com/journals/jama/article-abstract/2729375>, doi:10.1001/jama.2019.2413.
- Han, K., Pan, H., Gao, R., Yu, J., Yang, B., 2019. Multimodal 3d convolutional neural networks for classification of brain disease using structural mr and fdg-pet images, in: *International Conference of Pioneering Computer Scientists, Engineers and Educators*, Springer. pp. 658–668. URL: https://link.springer.com/chapter/10.1007/978-981-15-0118-0_51.
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., Larochelle, H., 2017. Brain tumor segmentation with deep neural networks. *Medical Image Analysis* 35, 18–31. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1361841516300330>, doi:<https://doi.org/10.1016/j.media.2016.05.004>.
- Hoffman, J., Tzeng, E., Park, T., Zhu, J.Y., Isola, P., Saenko, K., Efros, A., Darrell, T., 2018. Cycada: Cycle-consistent adversarial domain adaptation, in: *International conference on machine learning*, pp. 1989–1998. URL: <https://proceedings.mlr.press/v80/hoffman18a.html>.
- Hoffman, J., Wang, D., Yu, F., Darrell, T., 2016. Fcns in the wild: Pixel-level adversarial and constraint-based adaptation. arXiv preprint arXiv:1612.02649 URL: <https://arxiv.org/abs/1612.02649>.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708. URL: <https://arxiv.org/abs/1608.06993v5>.
- Huisman, A.G.M., 2009. Tumor-like lesions of the brain. *Cancer Imaging* 9, S10. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2797474/>.
- Ibtehaz, N., Rahman, M.S., 2020. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Networks* 121, 74–87. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0893608019302503>.
- Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134. URL: <https://arxiv.org/abs/1611.07004>.
- Jiang, J., Hu, Y.C., Tyagi, N., Zhang, P., Rimner, A., Mageras, G.S., Deasy, J.O., Veeraraghavan, H., 2018. Tumor-aware, adversarial domain adaptation from ct to mri for lung cancer segmentation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 777–785. URL: https://link.springer.com/chapter/10.1007/978-3-030-00934-2_86, doi:https://doi.org/10.1007/978-3-030-00934-2_86.
- Kamnitsas, K., Baumgartner, C., Ledig, C., Newcombe, V., Simpson, J., Kane, A., Menon, D., Nori, A., Criminisi, A., Rueckert, D., et al., 2017. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks, in: *International conference on information processing in medical imaging*, Springer. pp. 597–609. URL: https://link.springer.com/chapter/10.1007/978-3-319-59050-9_47, doi:https://doi.org/10.1007/978-3-319-59050-9_47.
- Karthik, R., Gupta, U., Jha, A., Rajalakshmi, R., Menaka, R., 2019. A deep supervised approach for ischemic lesion segmentation from multimodal mri using fully convolutional network. *Applied Soft Computing* 84, 105685. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1568494619304661>, doi:<https://doi.org/10.1016/j.asoc.2019.105685>.
- Kingma, D., Ba, J., 2014. Adam: A method for stochastic optimization. *Computer ence* URL: <https://arxiv.org/abs/1412.6980>.
- Korolev, S., Safiullin, A., Belyaev, M., Dodonova, Y., 2017. Residual and plain convolutional neural networks for 3d brain mri classification, in: *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*, IEEE. pp. 835–838. URL: <https://arxiv.org/abs/1701.06643>.
- Kothari, R.U., Brott, T., Broderick, J.P., Barsan, W.G., Sauerbeck, L.R., Zuccarello, M., Khoury, J., 1996. The abs of measuring intracerebral hemorrhage volumes. *Stroke* 27, 1304–1305. URL: <https://www.ahajournals.org/doi/full/10.1161/01.STR.27.8.1304>, doi:<https://doi.org/10.1161/01.STR.27.8.1304>.
- Lamon, P., Nourbakhsh, I., Jensen, B., Siegart, R., 2001. Deriving and matching image fingerprint sequences for mobile robot localization, in: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, IEEE. pp. 1609–1614. URL: <https://ieeexplore.ieee.org/document/932841>.
- Lei, Y., Dong, X., Tian, Z., Liu, Y., Tian, S., Wang, T., Jiang, X., Patel, P., Jani, A.B., Mao, H., et al., 2020. Ct prostate segmentation based on synthetic mri-aided deep attention fully convolution network. *Medical physics* 47, 530–540. URL: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13933>, doi:<https://doi.org/10.1002/mp.13933>.

- Lei, Y., Harms, J., Wang, T., Liu, Y., Shu, H.K., Jani, A.B., Curran, W.J., Mao, H., Liu, T., Yang, X., 2019. Mri-only based synthetic ct generation using dense cycle consistent generative adversarial networks. *Medical physics* 46, 3565–3581. URL: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13617>, doi:<https://doi.org/10.1002/mp.13617>.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440. URL: https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html.
- Luan, F., Paris, S., Shechtman, E., Bala, K., 2017. Deep photo style transfer, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4990–4998. URL: <https://ieeexplore.ieee.org/abstract/document/8100223>, doi:10.1109/CVPR.2017.740.
- Ma, J., Jiang, X., Fan, A., Jiang, J., Yan, J., 2021. Image matching from handcrafted to deep features: A survey. *International Journal of Computer Vision* 129, 23–79. URL: <https://link.springer.com/article/10.1007/s11263-020-01359-2>.
- Ma, W., Yu, S., Ma, K., Wang, J., Ding, X., Zheng, Y., 2019. Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer. pp. 769–778. URL: <https://arxiv.org/abs/2007.09337>.
- Maas, A.L., Hannun, A.Y., Ng, A.Y., 2013. Rectifier nonlinearities improve neural network acoustic models, in: *Proc. icml*, p. 3. URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.693.1422>, doi:10.21437/interspeech.2016-1230.
- Muschelli, J., Sweeney, E.M., Ullman, N.L., Vespa, P., Hanley, D.F., Crainiceanu, C.M., 2017. PItcHPERFeCT: Primary intracranial hemorrhage probability estimation using random forests on CT. *NeuroImage: Clinical* 14, 379–390. URL: <https://www.sciencedirect.com/science/article/pii/S2213158217300414>.
- O'donnell, M.J., Xavier, D., Liu, L., Zhang, H., Chin, S.L., Rao-Melacini, P., Rangarajan, S., Islam, S., Pais, P., McQueen, M.J., et al., 2010. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the interstroke study): a case-control study. *The Lancet* 376, 112–123. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0140673610608343>, doi:[https://doi.org/10.1016/S0140-6736\(10\)60834-3](https://doi.org/10.1016/S0140-6736(10)60834-3).
- Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al., 2018. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999* URL: <https://arxiv.org/abs/1804.03999>.
- de Oliveira Manoel, A.L., 2020. Surgery for spontaneous intracerebral hemorrhage. *Critical Care* 24, 1–19. URL: <https://link.springer.com/article/10.1186/s13054-020-2749-2>.
- Pan, S.J., Yang, Q., 2009. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 1345–1359. URL: <https://ieeexplore.ieee.org/abstract/document/5288526/>.
- Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016. Brain tumor segmentation using convolutional neural networks in mri images. *IEEE Transactions on Medical Imaging* 35, 1240–1251. URL: <https://ieeexplore.ieee.org/abstract/document/7426413>, doi:10.1109/TMI.2016.2538465.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer. pp. 234–241. URL: https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28, doi:https://doi.org/10.1007/978-3-319-24574-4_28.
- Roy, S., Carass, A., Jog, A., Prince, J.L., Lee, J., 2014. Mr to ct registration of brains using image synthesis, in: *Medical Imaging 2014: Image Processing*, International Society for Optics and Photonics. p. 903419. URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9034/903419/MR-to-CT-registration-of-brains-using-image-synthesis/10.1117/12.2043954.short?SSO=1>.
- Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition, in: *ICLR 2015 : International Conference on Learning Representations 2015*. URL: <https://academic.microsoft.com/paper/2962835968>.
- Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., Van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., et al., 2019. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv preprint arXiv:1902.09063* URL: <https://arxiv.org/abs/1902.09063>.
- Tsai, Y.H., Hung, W.C., Schulter, S., Sohn, K., Yang, M.H., Chandraker, M., 2018. Learning to adapt structured output space for semantic segmentation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7472–7481. URL: https://openaccess.thecvf.com/content_cvpr_2018/html/Tsai_Learning_to_Adapt_CVPR_2018_paper.html.
- Tzeng, E., Hoffman, J., Saenko, K., Darrell, T., 2017. Adversarial discriminative domain adaptation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7167–7176. URL: https://openaccess.thecvf.com/content_cvpr_2017/html/Tzeng_Adversarial_Discriminative_Domain_CVPR_2017_paper.html, doi:10.1109/CVPR.2017.316.
- Ulyanov, D., Vedaldi, A., Lempitsky, V., 2016. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022* URL: <https://arxiv.org/abs/1607.08022>.
- Wallis, D., Buvat, I., 2022. Clever hans effect found in a widely used brain tumour mri dataset. *Medical Image Analysis*, 102368 URL: <https://www.sciencedirect.com/science/article/pii/S1361841522000214>.
- Wang, S., Shen, Y., Chen, W., Xiao, T., Hu, J., 2017. Automatic recognition of mild cognitive impairment from mri images using expedited convolutional neural networks, in: *International Conference on Artificial Neural Networks*, Springer. pp. 373–380. URL: https://link.springer.com/chapter/10.1007/978-3-319-68600-4_43.
- Xiao, X., Lian, S., Luo, Z., Li, S., 2018. Weighted res-unet for high-quality retina vessel segmentation, in: *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, IEEE. pp. 327–331. URL: <https://ieeexplore.ieee.org/abstract/document/8589312>, doi:10.1109/ITME.2018.00080.
- Xie, G.S., Liu, J., Xiong, H., Shao, L., 2021. Scale-aware graph neural network for few-shot semantic segmentation, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5475–5484. URL: https://openaccess.thecvf.com/content/CVPR2021/papers/Xie_Scale-Aware_Graph_Neural_Network_for_Few-Shot_Semantic_Segmentation_CVPR_2021_paper.pdf.
- Xie, G.S., Zhang, X.Y., Yan, S., Liu, C.L., 2015. Hybrid cnn and dictionary-based models for scene recognition and domain adaptation. *IEEE Transactions on Circuits and Systems for Video Technology* 27, 1263–1274. URL: <https://arxiv.org/abs/1601.07977>.
- Zhang, Y., Wei, Y., Wu, Q., Zhao, P., Niu, S., Huang, J., Tan, M., 2020. Collaborative unsupervised domain adaptation for medical image diagnosis. *IEEE Transactions on Image Processing* 29, 7834–7844. URL: <https://ieeexplore.ieee.org/abstract/document/9142394>.
- Zhou, Y., He, X., Huang, L., Liu, L., Zhu, F., Cui, S., Shao, L., 2019. Collaborative learning of semi-supervised segmentation and classification for medical images, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2079–2088. URL: https://openaccess.thecvf.com/content_cvpr_2019/html/Zhou_Collaborative_Learning_of_Semi-Supervised_Segmentation_and_Classification_for_Medical_Images_CVPR_2019_paper.html.
- Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J., 2018. Unet++: A nested u-net architecture for medical image segmentation, in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer. pp. 3–11. URL: https://link.springer.com/chapter/10.1007/978-3-030-00889-5_1, doi:https://doi.org/10.1007/978-3-030-00889-5_1.
- Zhu, J.Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks, in: *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232. URL: https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html.

- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., He, Q., 2020. A comprehensive survey on transfer learning. *Proceedings of the IEEE* 109, 43–76. URL: <https://ieeexplore.ieee.org/abstract/document/9134370>.
- Zou, D., Zhu, Q., Yan, P., 2020. Unsupervised domain adaptation with dual-scheme fusion network for medical image segmentation., in: *IJCAI*, pp. 3291–3298. URL: <https://www.ijcai.org/proceedings/2020/0455.pdf>.
- Zou, F., Chen, Y., Song, J., Zhou, K., Yang, Y., Sebe, N., 2015. Compact image fingerprint via multiple kernel hashing. *IEEE transactions on multimedia* 17, 1006–1018. URL: <https://ieeexplore.ieee.org/document/7091952>.