A teacher–student framework with Fourier Transform augmentation for COVID-19 infection segmentation in CT images

Han Chen [a], Yifan Jiang [a], Hanseok Ko [a,*], Murray Loew [b]

[a] School of Electrical Engineering, Korea University, Seoul, South Korea
[b] Biomedical Engineering, George Washington University, Washington D.C., USA

ARTICLE INFO

ABSTRACT

Automatic segmentation of infected regions in computed tomography (CT) images is necessary for the initial diagnosis of COVID-19. Deep-learning-based methods have the potential to automate this task but require a large amount of data with pixel-level annotations. Training a deep network with annotated lung cancer CT images, which are easier to obtain, can alleviate this problem to some extent. However, this approach may suffer from a reduction in performance when applied to unseen COVID-19 images during the testing phase, caused by the difference in the image intensity and object region distribution between the training set and test set. In this paper, we proposed a novel unsupervised method for COVID-19 infection segmentation that aims to learn the domain-invariant features from lung cancer and COVID-19 images to improve the generalization ability of the segmentation network for use with COVID-19 CT images. First, to address the intensity difference, we proposed a novel data augmentation module based on Fourier Transform, which transfers the annotated lung cancer data into the style of COVID-19 image. Secondly, to reduce the distribution difference, we designed a teacher–student network to learn rotation-invariant features for segmentation. The experiments demonstrated that even without getting access to the annotations of the COVID-19 CT images during the training phase, the proposed network can achieve a state-of-the-art segmentation performance on COVID-19 infection.

1. Introduction

The pandemic caused by the novel coronavirus disease (COVID-19) that emerged at the start of 2020 has had significant worldwide medical and economic impacts [1,2]. A recent report by the World Health Organization found that there had been more than 198 million confirmed cases and more than 5 million deaths globally by January 2022 [3]. To diagnose COVID-19, real-time reverse transcription-polymerase chain reaction (RT-PCR) tests and radiological imaging techniques, such as computed tomography (CT) and X-rays, are widely used. In particular, CT imaging plays a critical role in the early diagnosis and evaluation of COVID-19 [4,5], with the segmentation of infected regions in CT scans providing essential information that can be used in the quantitative assessment of the progression of the disease [6,7].

Overall, segmentation approaches can be split into three groups: (1) Manual-based segmentation is defined as the delineation of the borders of anatomical regions that are conducted by experts (e.g., radiologists, pathologists) [8]; (2) Conventional methods-based segmentation is described as the assignment of labels to pixels or voxels by matching the prior known object model to the image data according to the radiomics features (hand-crafted or explicit features) [9]; (3) Deep-learning-based

segmentation as developed for automated feature extraction [10–12]. Segmentation of the lungs in COVID-19 cases involves outlining the contours of the anatomical structures of the lung or infection with computer-assisted techniques in CT or X-ray images. Due to the high variety in shape, size, boundary, type, and manifestation of regions of interest (ROI), the conventional algorithms failed to segment the infected region properly [8]. Some deep-learning-based automatic segmentation methods have been proposed for use in COVID-19 segmentation and achieved impressive results [13–15].

However, current COVID-19 infection segmentation approaches rely on large datasets with pixel-level annotations, which is time-consuming and laborious to collect due to privacy issues and the lack of experts [16–18]. In contrast, as one of the most common cancers worldwide, the research on lung cancer is more widely, and the labeled lung cancer CT data is more accessible [19]. Therefore, it is possible to utilize publicly available lung cancer databases to promote the performance of a deep network for the segmentation of COVID-19 infection. For example, Jin et al. [20] developed a system using a large dataset from the Lung Image Database Consortium and Image Database

(a) Lung Cancer CT Image   (b) COVID-19 CT Image
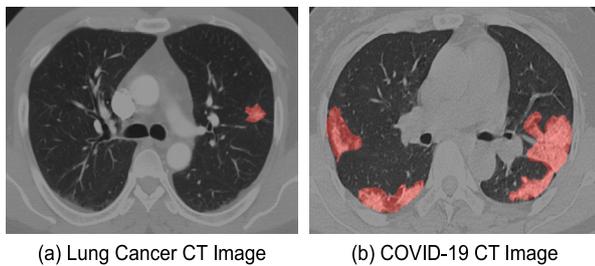
Fig. 1. Comparison of the pulmonary nodule and COVID-19 infection in CT images. (a) shows a typical nodule of lung cancer and (b) illustrates the infection of a COVID-19 confirmed case. The nodule and infection are marked with red.

Resource Initiative (LIDC-IDRI) [21] to achieve multi-class classification diagnoses. Wang et al. [22] revealed that transferring knowledge of lung cancer datasets can boost the segmentation performance for COVID-19 infection. They proved that a robust pre-training model can be obtained by training a segmentation model with lung cancer data with pulmonary nodule, which benefits subsequent supervised learning on a small COVID-19 dataset.

Even though COVID-19 infection and pulmonary nodule exhibit similar manifestations, few studies have directly used lung cancer datasets to train the deep network for COVID-19 infection segmentation. Models only trained with lung cancer CT images will suffer a significant performance drop when tested on COVID-19 CT images due to the domain shift between the two datasets. The difference between pulmonary nodules and COVID-19 infection in CT scans is presented in Fig. 1 and can be categorized as follows: (1) In terms of distribution, COVID-19 presents as a bilateral, patchy infection, while early-stage lung cancer is unilateral and oval in shape [23]; (2) While there is also a clear difference in intensity due to the use of different scanners, scanning protocols, and subject populations [24].

In this paper, we consider COVID-19 infection segmentation in the context of the wide availability of lung cancer CT images with annotations, the limited availability of unlabeled COVID-19 CT images, and the difference between these two domains. Our motivation is that the features learned from pulmonary nodules in lung cancer CT can be used for the segmentation of COVID-19 infection [22], making it possible to construct an unsupervised COVID-19 infection segmentation method. For our segmentation network, we designed a novel COVID-19 style guided Fourier Transform-based data augmentation method (CGFT-DA) and a training scheme to align these two datasets. In order to address the intensity difference, we transferred the lung cancer CT images into the style of COVID-19 CT images with the designed CGFT-DA module, which replaces the low-level frequency information of the lung cancer CT images with that of COVID-19 CT images. The output of the CGFT-DA module will keep the same semantic information with the lung cancer image but in the COVID-19 image style. Since the lung cancer CT images are labeled at pixel-level, the transferred images and corresponding annotations can be used to train the base segmentation network. To overcome the distribution difference, we introduced a teacher–student learning paradigm to achieve robust features learning. We treated our base segmentation network as the *student* network and introduced another teacher network, and then imposed the same elastic transformation on the input to the student network and the output of the teacher network, respectively. The output predictions of these two networks are forced to be consistent. We validated the effectiveness of our proposed method with public COVID-19 CT images. Experimentally, it outperformed various competing state-of-the-art approaches.

The contributions of our work can be summarized as follows:

- We proposed a novel unsupervised COVID-19 infection segmentation method. To the best of our knowledge, our method is the first attempt to utilize the lung cancer data for the COVID-19 infection segmentation task.

- We proposed an effective data augmentation module to overcome the intensity difference between the lung cancer CT data and COVID-19 CT data via the Fourier Transform and its inverse.
- We built a framework that collaborates with a teacher–student network based on transformation-consistency learning, which alleviates the distribution difference and enables the network to learn the robust features of COVID-19 infection in CT images.
- The experimental results showed that our proposed method can achieve state-of-the-art performance for COVID-19 infection segmentation. We also provided a comprehensive analysis of our method.

2. Related works

COVID-19 infection segmentation. Deep-learning-based segmentation methods have played an essential role in diagnosing COVID-19 [25–29]. Some studies explored to construct new networks suitable for COVID-19 CT segmentation task [30–32]. For example, Ouyang et al. [30] developed an online module with a 3D convolutional network (CNN) for better localization of COVID-19 infected areas. Other works utilized attention mechanism to learn rich contextual relationships for better feature representations [33–35]. Zhou et al. [33] proposed to use spatial and channel attention to improve the representation ability of the segmentation network. To address the issue of scarcity of well-labeled data, solutions based on weak supervised, semi-supervised learning and multi-task learning were proposed in the literature [36–38]. Fan et al. [36] presented a semi-supervised segmentation method based on a random selection propagation strategy, which required only a few labeled images and primarily utilizes unlabeled data. However, current techniques still rely more or less on labeled COVID-19 data and cannot handle the more challenging unsupervised situation. Meanwhile, few studies consider using lung cancer data which is easier to access, to boost the segmentation performance of COVID-19 infection.

Semi-supervised learning for medical segmentation. Deep-learning-based methods usually require a large amount of labeled data, however, it is expensive and time-consuming to annotate data for medical segmentation tasks. Semi-supervised learning gained significant attention as a way to utilize the unlabeled data [39]. Some studies are based on self-training, which consists of pre-training the segmentation network using unlabeled medical images followed by fine-tuning the pre-trained model using the target labeled dataset [40–42]. Other works are based on pseudo annotations, they first assigned pseudo annotations to unlabeled data and then trained the segmentation model using both the labeled and pseudo labeled data [43–45]. For example, Luo et al. [43] proposed to utilize cross training and take the prediction of a network as pseudo-label for other networks. Another strategy is training a model with both labeled and unlabeled data, where the unlabeled data is used to generate a supervision signal through an unsupervised loss function [46–48]. Zhang et al. [46] proposed an adversarial network for biomedical image segmentation, which is composed of an evaluation network to assess the segmentation performance and encourage the network to attain consistently good results on both annotated and unannotated images.

Teacher–student learning. The concept of teacher–student learning is based on distilling the knowledge of an ensemble of deep networks into a single network [49,50]. It employs unlabeled data to produce consistent predictions under perturbations and has been explored in segmentation tasks [51–53]. Yu et al. [51] proposed training a teacher–student model in which the teacher model produced noisy labels and corresponding uncertainty maps, while the student model was trained using the labels from the teacher model in consideration of label uncertainty. Xu et al. [54] proposed a self-ensembling attention network consisting of a student model that acts as a base network and a teacher model serving as the ensembling network, the former learns from the output of the latter and tends to produce accurate predictions. Choi
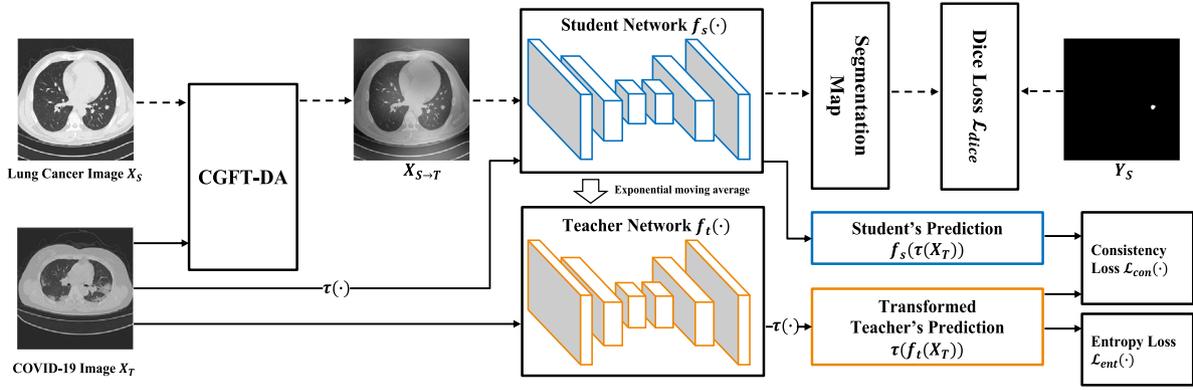
Fig. 2. Overview of the proposed method. It consists of a CGFT-DA module, a student network, and a teacher network. The dashed lines and solid lines represent the data flow for the lung cancer images and COVID-19 images, respectively. The student network is trained by the weighted combination of dice loss, consistency loss, and entropy loss. The weights of the teacher network are the exponential moving average of that of the student network. $\tau(\cdot)$ represents an elastic transformation operation with the same parameters in every iteration.

et al. [55] utilized GAN to transfer the labeled image into the same style as the unlabeled image and then introduced a teacher network to enhance its performance in the segmentation of the unlabeled data.

To compensate for the data shortage issue for the COVID-19 infection segmentation task, training one semi-supervised network with large-scale lung cancer CT data and unlabeled COVID-19 CT data could be a possible solution. However, the current semi-supervised techniques mainly focus on utilizing the unlabeled data, which shares similar distribution with the labeled data due to coming from the same source. Therefore they cannot address the more challenging situation of significant differences between these two datasets. To solve the above problem, we introduced our method. Specifically, we proposed to transfer the lung cancer data into the style of COVID-19 CT images for data augmentation to overcome the intensity difference. By constructing a teacher–student network with consistency learning and entropy minimization, our network can learn the features that are robust to elastic transformations, at the same time, the distribution difference gets alleviated effectively.

3. Proposed method

In this section, we provide an overview of our proposed method. We then illustrate our Fourier Transform-based data augmentation method. Finally, we describe our teacher–student training strategy and optimization objection.

3.1. Overview

Fig. 2 overviews our unsupervised COVID-19 infection segmentation network. The labeled lung cancer images and unlabeled COVID-19 CT images are denoted as $\{X_S, Y_S\} \sim D_S$ and $\{X_T\} \sim D_T$, respectively. We first send X_S and X_T to CGFT-DA module and obtain $\{X_{S \to T}, Y_S\} \sim D_{S \to T}$ which retains the semantic information for lung cancer but in COVID-19 style. A teacher and a student network are employed in our framework, with the former acting as an ensemble network and the latter acting as the base segmentation network. We send $X_{S \to T}$ into the student network and obtain the pixel-wise segmentation prediction. X_T will be sent to both the student and teacher networks, we then apply the same elastic transformation to the input of the student network and to the output of the teacher network, the two-stream outputs are aligned by the consistency loss. Back-propagation only occurs for the student network, and the weights of the teacher network are updated using the exponential moving average (EMA) of the student network.
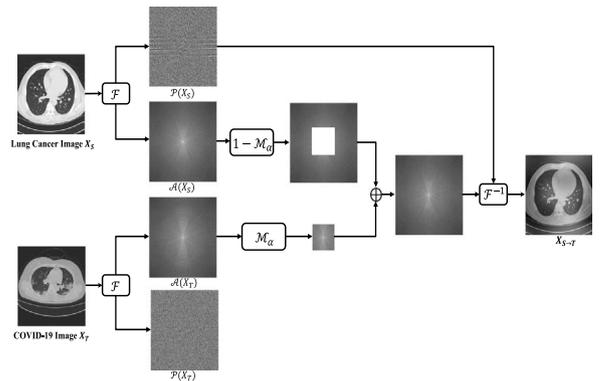
Fig. 3. The overview of the CGFT-DA module. The Fourier Transform is first conducted on the two inputs, and part of the amplitude spectrum from the lung cancer image is replaced with that from the COVID-19 image. After inverse Fourier Transform, the output image will keep the semantic information of the lung cancer image but with COVID-19 style.

3.2. COVID-19 style guided and fourier transform-based data augmentation

To address the intensity difference between the lung cancer and COVID-19 CT images, we proposed a COVID-19 style guided Fourier Transform-based data augmentation module (CGFT-DA), as shown in Fig. 3. We employ the Fourier Transform and replace the low-frequency spectrum information of the lung cancer images with that of the COVID-19 images. In this way, we make it possible to disentangle the low-level distribution (i.e., the intensity information) from the high-level semantic information (i.e., the object content) of an image and transfer the former to another image.

Specifically, given an image $X \in \mathbb{R}^{H \times W \times C}$ ($C = 1$ for a single-channel input), we can calculate its Fourier Transform using the FFT algorithm as,

$$\mathcal{F}(X)(u, v) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X(h, w) e^{-j2\pi(\frac{h}{H}u + \frac{w}{W}v)} \tag{1}$$

We further decompose $\mathcal{F}(X)$ into an amplitude spectrum $\mathcal{A}(X) \in \mathbb{R}^{H \times W \times C}$ and a phase spectrum $\mathcal{P}(X) \in \mathbb{R}^{H \times W \times C}$, which represent the intensity distribution and semantic content of the image, respectively. We then denote a binary mask $\mathcal{M}_\alpha = \mathbb{1}_{(h,w) \in [-\alpha H : \alpha H, -\alpha W : \alpha W]}$, which controls the scale of the amplitude spectrum to be replaced. Given $\mathcal{A}(X_S)$ from the lung cancer image and $\mathcal{A}(X_T)$ from the COVID-19 image, we apply the mask $\mathcal{M}_\alpha$ to them and generate a new amplitude spectrum:

$$\mathcal{A}(X_{S \to T}) = \mathcal{M}_\alpha \mathcal{A}(X_T) + (1 - \mathcal{M}_\alpha) \mathcal{A}(X_S) \tag{2}$$

After obtaining the transferred amplitude spectrum $\mathcal{A}(X_{S\rightarrow T})$, we combine it with the phase spectrum $\mathcal{P}(X_S)$ for lung cancer image X_S and conduct the inverse Fourier Transform $\mathcal{F}^{-1}$ to generate transferred image $X_{S\rightarrow T}$.

$$X_{S\rightarrow T} = \mathcal{F}^{-1}(\mathcal{A}(X_{S\rightarrow T}), \mathcal{P}(X_S)) \tag{3}$$

where the semantic content of $X_{S\rightarrow T}$ is the same as X_S, but follows the intensity distribution of X_T. In this way, we obtain the transferred lung cancer data $\{X_{S\rightarrow T}, Y_S\}$ with a similar intensity to that of COVID-19 images; this transferred data is utilized for further training and can effectively alleviate the intensity distribution between the two datasets.

3.3. Teacher–student network and entropy minimization

Even though the transferred data $X_{S\rightarrow T}$ have a similar style to X_T, a difference in distribution remains. To account for this distribution difference between the pulmonary nodules and COVID-19 infection, we built a teacher–student network with a transformation-consistency learning strategy. The teacher and student networks follow the same U-Net architecture [56] but are updated in different ways.

Specifically, the transferred lung cancer images $X_{S\rightarrow T}$ is only sent to the student network to output the pixel-wise segmentation result $\hat{Y}_S$, which is then utilized to calculate the following dice loss [57]:

$$\mathcal{L}_{dice} = \frac{2 * \sum_{n=1}^{N} p_n g_n}{\sum_{n=1}^{N} p_n + \sum_{n=1}^{N} g_n + \epsilon} \tag{4}$$

where g_n and p_n represent the ground truth for the input and the predicted probabilistic map, respectively, with a background class probability of $1-p_n$. The term ϵ is used to avoid being divided by zero.

Updating the base student network with the dice loss in Eq. (4) leads to good segmentation performance for lung cancer images. However, when testing on the COVID-19 images, the performance will be significantly lower due to the distribution difference. To resolve this problem, we introduced a teacher network and utilize the unlabeled COVID-19 CT images to guide the student network to learn the robust features. As shown in Fig. 2, we first impose an elastic transformation $\tau(\cdot)$ on X_T, then the student network takes the transformed $\tau(X_T)$ as input and produces the predicted segmentation map. For the teacher network, we apply the same transformation $\tau(\cdot)$ on its prediction map. We then define the consistency loss to force the predictions for COVID-19 infection from these two networks to be consistent:

$$\mathcal{L}_{con} = \sum_{i\in I} \|f_s(\tau(X_T))_i - \tau(f_t(X_T))_i\|^2 \tag{5}$$

To more firmly bridge the gap between the lung cancer images and the COVID-19 images, we introduced entropy minimization to force the model's decision boundary toward a high prediction certainty for the teacher network's prediction from the COVID-19 input. Given a COVID-19 CT image X_T, the entropy loss is calculated as follows:

$$\mathcal{L}_{ent} = - \sum_{i\in I} f_t(X_T)_i log(f_t(X_T))_i \tag{6}$$

We update the student network with a combination of the dice loss, consistency loss and entropy loss. The optimization objection can thus be formulated as follows:

$$\mathcal{L}_{student} = \mathcal{L}_{dice} + \lambda\mathcal{L}_{con} + \mathcal{L}_{ent} \tag{7}$$

where λ is the hyper-parameter acting as a weight for the consistency loss. Unlike the student network, the teacher network does not participate in the back-propagation and is updated via the exponential moving average (EMA) of the weights for the current student network at each step,

$$\theta_{t,j} = \beta\theta_{t,j-1} + (1-\beta)\theta_{s,j} \tag{8}$$

where $\theta_{t,j}$ and $\theta_{s,j}$ represent the weights for the teacher and student networks at training step j, respectively, and β is a hyper-parameter for exponential moving average decay.

Algorithm 1: Model Training

Input	: CGFT-DA processed dataset $D_{S\rightarrow T}$
	COVID-19 dataset D_T
Output	: Trained parameters θ^*
Parameters:	Two coefficients λ and β
	An elastic transform function τ
	An model forward function f

1 *Initialize network parameters θ^* randomly*;
2 **for** $j = 1$ *to* N **do**
3   $X_{S\rightarrow T}, Y_S \sim D_{S\rightarrow T}$;
4   $X_T \sim D_T$;
5   $\hat{Y}_S \leftarrow f_S(X_{S\rightarrow T})$;
6   *Compute Dice Loss* $\mathcal{L}_{dice} \leftarrow \frac{2*\sum_{n=1}^{N} p_n g_n}{\sum_{n=1}^{N} p_n + \sum_{n=1}^{N} g_n + \epsilon}$;
7   *Compute Consistency Loss*
     $\mathcal{L}_{con} \leftarrow \sum_{i\in I} \|f_s(\tau(X_T))_i - \tau(f_t(X_T))_i\|^2$;
8   *Compute Entropy Loss* $\mathcal{L}_{ent} \leftarrow - \sum_{i\in I} f_t(X_T)_i log(f_t(X_T))_i$;
9   $\mathcal{L}_{student} = \mathcal{L}_{dice} + \lambda\mathcal{L}_{con} + \mathcal{L}_{ent}$;
10   *Update θ_s by backpropagation*;
11   *Update θ_t using EMA* $\theta_{t,i} = \beta\theta_{t,j-1} + (1-\beta)\theta_{s,j}$
12 **end**

The ensemble of the teacher and student network makes it possible to train the network with the unlabeled COVID-19 CT images. Moreover, through the regularization of the consistency loss, the student network can learn from the teacher network output. The networks are then regularized to be transformation-consistent thus increasing the generalization capacity and robustness to the distribution difference between the lung cancer images and the unlabeled COVID-19 images.

4. Experiments

4.1. Experimental settings

Datasets. The lung cancer data comes from the LIDC-IDRI lung cancer dataset [21], which is currently the largest CT dataset for pulmonary nodule detection. It provides a large number of lung CT images that share similarities with COVID-19 CT images. The LIDC-IDRI dataset contains 1018 cases, each of which includes images from a clinical thoracic CT scan along with pixel-wise annotations from experienced radiologists. The location information of every pulmonary nodule is recorded in an associated XML file for each case. The COVID-19 CT data come from a collection which is made by the Italian Society of Medical and International Radiology [58]. This dataset includes the original CT images and COVID-19 infection masks. It contains 26 CT volumes from confirmed COVID-19 patients, and each volume contains ~200 slices. The data processing for the above two datasets is detailed as follows,

LIDC-IDRI lung cancer dataset. The data pre-processing steps are as follows. (1) 3D volumes to 2D slices: We reformatted all the 3D volumes into 2D slices with a size of 512×512. (2) Exclude scans with small nodules: To capture the presence of the nodules that might be meaningful for COVID-19 infection segmentation without decreasing the network's performance by a preponderance of smaller nodules, we excluded the scans where the pixel number for every nodule is less than 200 for robust training. (3) Generate lung nodule mask: We generated ground-truth nodule masks for scans selected by step(1) based on the patient's XML file. (4) CT image denoising: all CT scans are transformed to gray images on a Hounsfield unit (HU) scale [−600, 1500] to remove noise in CT images. The above pre-processing leaves 2438 slices for lung cancer CT images with annotations.

COVID-19 dataset. The processing of the COVID-19 dataset is similar to that of the LIDC-IDRI dataset but without step (2). After processing,

Table 1
Organization of the dataset for training and test. Note that the ground-truth annotations of the COVID-19 training set are not accessible during training.

Dataset	Patients	Slices count
Training set (Lung cancer)	236	2438
Training set (COVID-19)	18	1156
Test set (COVID-19)	8	460

there are a total of 1616 2D COVID-19 CT slices. We employed 70% of these slices as the unlabeled data for training, while the remaining 30% are used to test segmentation performance. We followed the patient-level split rule when separating the COVID-19 data into a training set and a test set. Please see the detailed datasets organization in Table 1.

Implementation details. The batch size was set to 1. We used an Adam optimizer with the initial learning rate of 6e-4 and weight decay of 0.0005. The hyper-parameters were set at $\alpha = 0.005$ and $\beta = 0.99$. The consistency weight was updated using a sigmoid ramp-up of $\lambda = 1.5e^{-5(1-p)^2}$, where p is the progress of the training epochs normalized to a range of 0 to 1.

The libraries involved in this work include PyTorch 1.7.1 [59] for network construction and Fourier Transform, Medicaltorch 0.2 [60] for elastic transformation, Seaborn 0.11.2 for boxplot [61], and Scikit-learn 0.24.2 [62] for feature visualization, etc. Our network is implemented on an NVIDIA RTX 2080Ti and an Intel(R) Core i7-9700K CPU.

The structure of the student and teacher networks follows U-Net architecture [56], which is implemented by the built-in function of PyTorch. The trained student network is taken as the segmentation network during the test phase and outputs the segmentation results for the input COVID-19 CT images.

Evaluation metrics. For quantitative evaluation, we adapted the three most commonly used metrics in medical imaging analysis: the dice similarity coefficient (Dice) [63], sensitivity (Sen) [64], and specificity (Spe) [65]. The dice similarity coefficient is an overlap index that indicates the similarity between the prediction and the ground truth. Sensitivity and specificity are two statistical measures for the performance of medical image segmentation tasks. The former measures the percentage of actual positive pixels correctly predicted to be positive, while the latter measures the proportion of actual negative pixels correctly predicted to be negative. These metrics are defined as follows:

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{9}$$

$$Sen = \frac{TP}{TP + FN} \tag{10}$$

$$Spe = \frac{TN}{TN + FP} \tag{11}$$

where TP, FP, TN, and FN represent the number of true positive, false positive, true negative, and false negative pixels in the prediction, respectively.

4.2. Experimental results

In this section, we compared the COVID-19 infection segmentation performance of the proposed method with that of the Source-only method (U-Net trained with the lung cancer data in a supervised manner) and the state-of-the-art semi-supervised methods: UA-MT [51], DAN [46] and CTCT [43] for medical segmentation; CCT [45], CPS [44], FDA [41], MinEnt [47], and AdvEnt [47] for cityscapes segmentation; MT [49], DCT [40] and ICT [48] originally designed for classification task, we adapted these semi-supervised classification methods to segmentation task for comparison.

Quantitative results. We evaluated the segmentation performance of our proposed method through comparison with the above state-of-the-art semi-supervised methods. We also tested the robustness of the

Table 2
Comparison between our proposed method and existing semi-supervised segmentation methods on the COVID-19 CT dataset. The highest evaluation score is marked in bold. ↑ indicates that a higher number is better.

Method	Dice (%) ↑	Sen (%) ↑	Spe (%) ↑
Source-only	40.06 ± 8.26	33.68 ± 8.55	99.86 ± 0.03
UA-MT [51]	40.95 ± 5.96	33.40 ± 6.49	99.80 ± 0.04
DAN [46]	42.87 ± 3.42	32.81 ± 3.81	99.85 ± 0.03
CTCT [43]	44.53 ± 6.11	43.73 ± 7.05	99.21 ± 0.19
CCT [45]	45.14 ± 4.99	40.35 ± 6.89	99.72 ± 0.07
CPS [44]	45.26 ± 5.85	36.63 ± 6.89	99.82 ± 0.04
FDA [41]	41.55 ± 8.75	38.28 ± 7.33	**99.96 ± 0.01**
MinEnt [47]	43.86 ± 6.43	40.65 ± 6.06	99.95 ± 0.02
AdvEnt [47]	44.11 ± 7.35	40.76 ± 6.54	99.95 ± 0.02
MT [49]	41.49 ± 5.28	30.15 ± 5.20	99.90 ± 0.02
DCT [40]	44.42 ± 5.34	34.23 ± 5.18	99.85 ± 0.04
ICT [48]	47.01 ± 3.97	37.74 ± 4.97	99.84 ± 0.04
Ours	**49.51 ± 6.48**	**44.70 ± 9.38**	99.75 ± 0.06
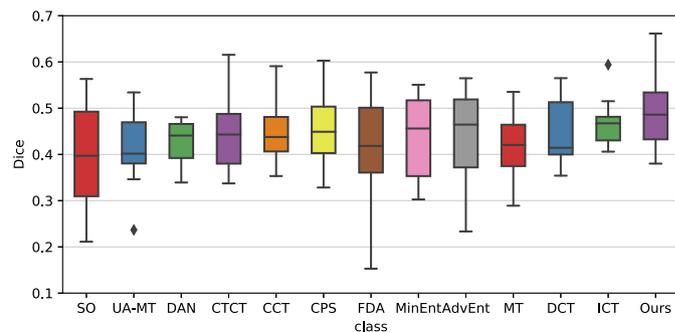
Fig. 4. Boxplot for the distribution of Dice similarity coefficient results on COVID-19 infection segmentation task, with the maximum/minimum results shown as upper/lower extreme, 95% confidence interval as the color box, and outlier as dot.

proposed method by visualizing the distribution of the results for the dice metric. For each experiment, we excluded the test set from the training of these deep learning-based models. After training, the test set was used for the segmentation test.

Table 2 shows the quantitative results for each method, reported as the mean± error interval (calculated based on a 95% confidence interval) of the Dice similarity coefficient, sensitivity, and specificity. Generally, more accurate segmentation results are associated with higher scores for the above metrics. Our proposed method outperformed the compared methods across most metrics. Regarding dice, our method produced a 9.54% improvement over the Source-only method, which confirms the effectiveness of our augmentation process and consistency learning. It also signifies that our approach can achieve more similarity between the predicted COVID-19 infection area and the ground truth infected area. Our method outperformed other methods on sensitivity (also known as the "true positive" rate), demonstrating that our network can identify the pixels that belong to the COVID-19 infected area more accurately. Compared with the UA-MT [51], and MT [49] which also adapted teacher–student framework, our proposed method utilized elastic transformation for consistency learning, which benefits improving the robustness of our model for the shape difference between pulmonary nodule and COVID-19 infection. Our method performed competitive results on specificity (also known as the "true negative" rate), proving that our proposed model can correctly rule out pixels that do not belong to the infected area and will not generate many false-positive results.

To investigate the robustness of our proposed method, we visualized the dice similarity coefficient results for all COVID-19 test slices of our method and the compared methods by boxplot. As shown in Fig. 4, FDA [41] and AdvEnt [47] showed large range for extreme value, representing these methods tend to perform unstable predicted segmentation
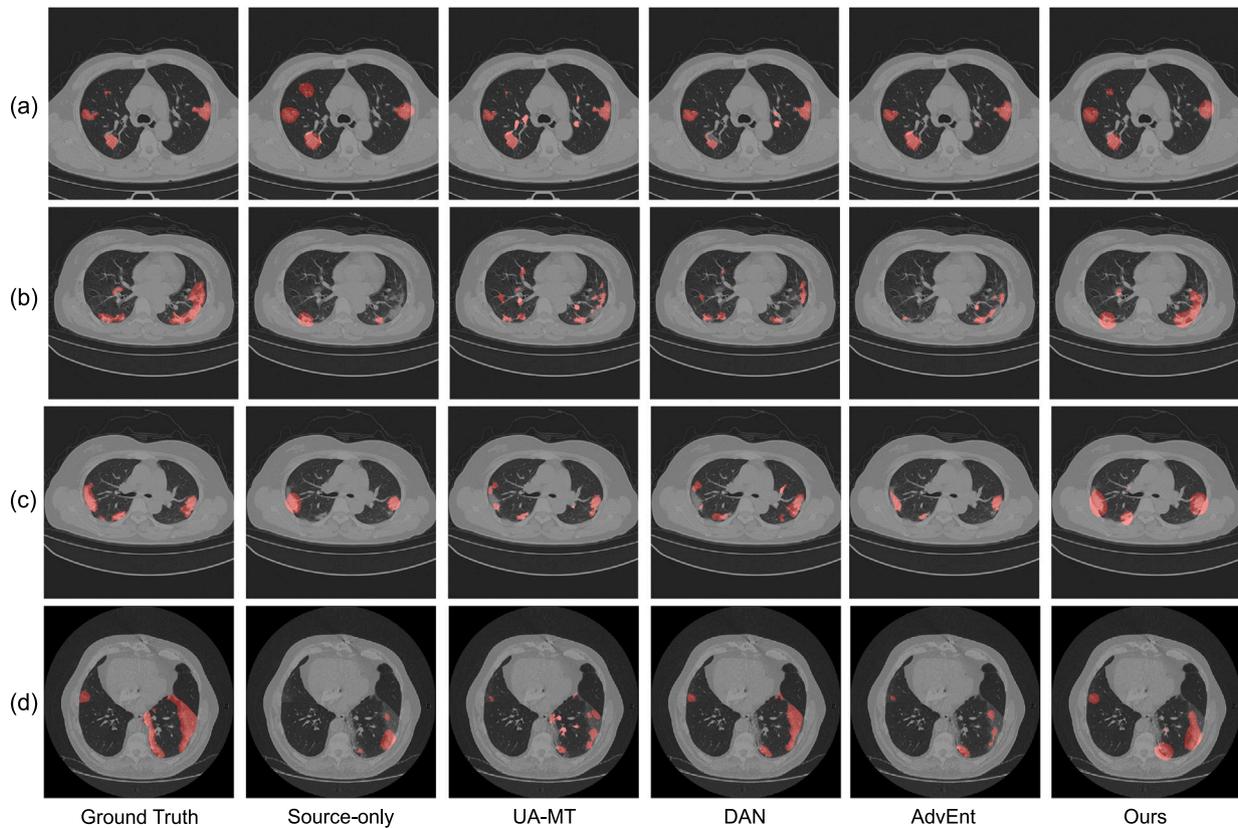
Fig. 5. Qualitative results for the segmentation task. The first column presents the input COVID-19 CT images with the ground truth marked in red, while columns 2 to 6 are the segmentation results of the Source-only method, UA-MT [51], DAN [46], AdvEnt [47], and our proposed method.

results on the different test slices. There are outliers (too low/high) appear in results of UA-MT [51] and ICT [48], which is also unacceptable for the real clinical application. In contrast, our proposed method yielded more stable and better results than other methods. It is mainly because the proposed CGFT-DA module can alleviate the appearance difference between lung cancer and COVID-19 images. Furthermore, our teacher–student framework enables the network to learn the robust features from unlabeled COVID-19 CT images gradually.

Qualitative results. To demonstrate the effectiveness of our proposed method from the qualitative perspective, we visualized the segmentation results of our method and some compared methods in Fig. 5. As shown in cases (a) to (d), the predicted infected area generated with our proposed method shows the best resemblance to the ground truth. The Source-only method can distinguish small infected regions but suffers from poor generalization for large-scale infection, such as case (b) because it is only trained with the LIDC-IDRI lung cancer dataset, in which the pulmonary nodules are relatively small and oval. UA-MT [51] showed some mis-segmentation in case (a) and (b). DAN [46] produced incomplete segmentation when handle with case (b). AdvEnt [47] also suffered from the poor generalization for the large-area COVID-19 infection in case (b) and (d). Overall, our proposed method performed better than the other methods and is consistently closer to the ground-truth COVID-19 infected region, demonstrating that our method effectively utilizes the unlabeled COVID-19 images and improves segmentation performance.

4.3. Ablation study

This section describes the ablation study. The purpose is to assess the importance of each component in our network. Moreover, we visualized the feature distributions as learned by the proposed method and the Source-only method to prove that our proposed method could

Table 3
Ablation study of contribution of each component in our proposed network. The highest evaluation score is marked in bold. ↑ indicates that a higher number is better.

Method	Dice (%) ↑	Sen (%) ↑	Spe (%) ↑
Source-only	40.06 ± 8.26	33.68 ± 8.55	99.86 ± 0.03
w/o CGFT-DA	48.40 ± 6.82	43.59 ± 11.33	99.63 ± 0.11
w/o con.	46.94 ± 7.70	39.13 ± 8.56	**99.89 ± 0.04**
w/o ent.	47.98 ± 7.07	44.45 ± 9.92	99.72 ± 0.05
Ours	**49.51 ± 6.48**	**44.70 ± 9.38**	99.75 ± 0.06

effectively align the lung cancer data and COVID-19 data in feature space.

Contribution of each proposed component. We validated the effect of each component of our network by analyzing the performance of the following setups: (1) Source-only, which includes only the student network and is trained with the lung cancer data in a supervised manner; (2) w/o CGFT-DA, in which the data augmentation process that is essential for overcoming the intensity difference is removed; (3) w/o con., in which the teacher network is removed, corresponding to $\lambda = 0$ and the consistency loss not being used to update the student network; (4) w/o ent., in which the entropy minimization is not calculated, which corresponds to $\lambda = 0$. As shown in Table 3, the exclusion of any of the components, especially the teacher network, leads to a drop in performance, thus confirming that these components play a vital role in the performance of our proposed method. In particular, data augmentation lays a foundation for alleviating the intensity difference, while the teacher–student training scheme allows robust transformation-invariant features to be effectively exploited.

Visualization of feature distributions via t-SNE. To verify that our proposed method can effectively overcome the gap between lung cancer CT data and COVID-19 CT data and learn the domain-invariant
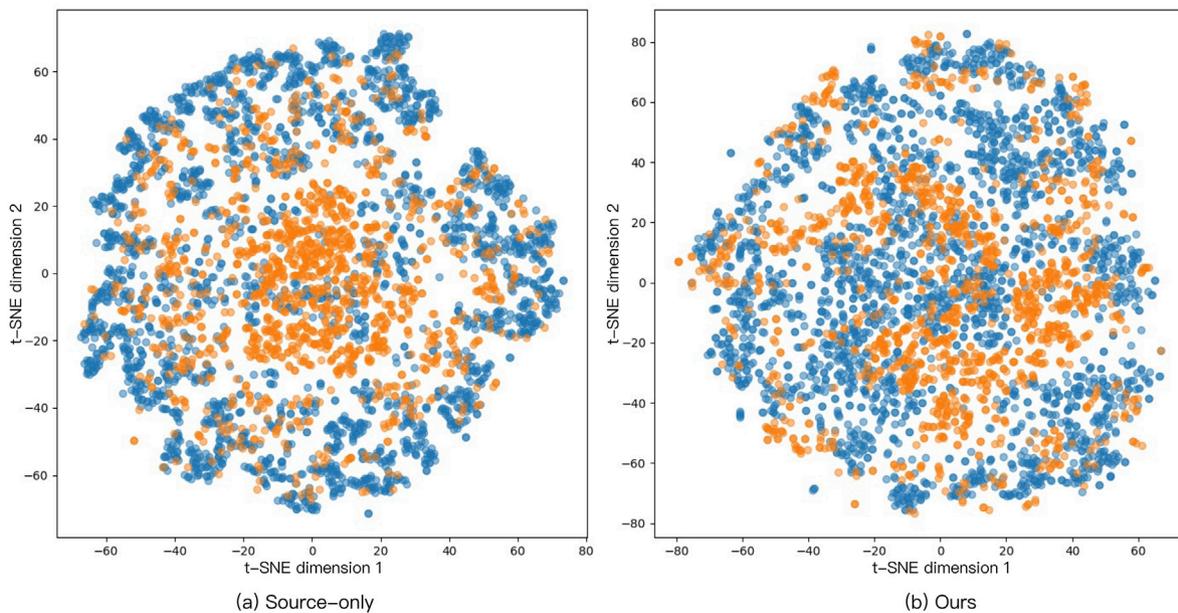
Fig. 6. Visualization of the feature distributions via t-SNE for two different datasets. The figure shows t-SNE visualization of the high-dimension feature extracted from (a) Source-only model trained on lung cancer dataset and (b) Our proposed model which incorporates lung cancer data and COVID-19 data for training. The blue dots correspond to the lung cancer samples, while the yellow dots correspond to the COVID-19 samples. In all cases, the Fourier Transform-based augmentation strategy and teacher–student training scheme in our method make the two feature distributions much closer.

features, we visualized the feature representations using t-SNE [66]. We input the lung cancer and COVID-19 CT images to the trained Source-only model and our proposed network, respectively, then visualize the feature maps for these two data groups in low dimensions. As shown in Fig. 6(a), for the Source-only model, the feature distributions of lung cancer images are separated from those of the COVID-19 images because supervision only occurs with the lung cancer data. Thus, the learned features from the Source-only cannot improve the segmentation performance for COVID-19 images. In contrast, as demonstrated in Fig. 6(b), our proposed method projects the feature distributions of the two datasets into an overlapping space, illustrating the effectiveness of our scheme in aligning the two distributions. The visualization results indicate that our proposed network can learn robust domain-invariant features.

5. Discussion

There are several limitations of the proposed method. First, the number of patients with COVID-19 is relatively few. Since this work took place at the early stage of the outbreak of this pandemic, we could only acquire a small set of data from one institution for training our network, so the method has not shown its most potential. In the future, we will focus on training the segmentation network with more sufficient COVID-19 data, including other modalities such as CT and MRI to improve the segmentation performance further.

Second, this work only involved the segmented COVID-19 infection in the lung. Theoretically, this method can be extended to other organs from CT images. For example, we can adapt our work for diagnosing primary hepatic tuberculosis that shares similarities with common hepatocellular carcinoma. Detecting and segmenting other organs from CT images using the proposed strategy would also be future work.

6. Conclusion

In this paper, we proposed a novel teacher–student-based framework for unsupervised COVID-19 infection segmentation in CT images. We attempted to address a challenging situation where there is no annotation for COVID-19 CT images, but the annotation for lung CT

with pulmonary nodules is available. Given the differences between the pulmonary nodules with the COVID-19 infection, we introduced a Fourier Transform-based augmentation method to alleviate the intensity difference. We further constructed a teacher–student network that utilized the consistency loss and entropy loss to allow the proposed network to learn the robust features, thus overcoming the distribution difference to some extent. Experiments on the COVID-19 CT dataset demonstrated that the proposed method could produce a competitive performance compared with the state-of-the-art methods. In conclusion, this technique could provide accurate segmentation for treatment planning without requiring COVID-19 pixel-wise annotation acquisition, greatly facilitating the routine clinical workflow. In the future, we plan to extend our work to a multimodal framework, which learns shared representations by training with datasets from different modalities (such as MRI and CT), thereby further addressing the data shortage problem and improving segmentation quality.

CRediT authorship contribution statement

Han Chen: Conceptualization, Methodology, Software, Writing – original draft. **Yifan Jiang:** Investigation, Data curation. **Hanseok Ko:** Supervision, Writing – review & editing. **Murray Loew:** Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hanseok Ko reports financial support was provided by The Government-Wide R&D Fund Project for Infectious Disease Research (GFID), Republic of Korea.

Data availability

Data will be made available on request.

References

[1] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395 (10223) (2020) 470–473.

[2] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, D. Shen, Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19, IEEE Rev. Biomed. Eng. 14 (2020) 4–15.

[3] Who Coronavirus (COVID-19) Dashboard, https://covid19.who.int/.

[4] J. Lei, J. Li, X. Li, X. Qi, CT imaging of the 2019 novel coronavirus (2019-nCoV) pneumonia, Radiology 295 (1) (2020) 18.

[5] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, et al., Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT, Radiology (2020).

[6] L. Huang, R. Han, T. Ai, P. Yu, H. Kang, Q. Tao, L. Xia, Serial quantitative chest CT assessment of COVID-19: a deep learning approach, Radiol.: Cardiothorac. Imaging 2 (2) (2020) e200075.

[7] Y. Cao, Z. Xu, J. Feng, C. Jin, X. Han, H. Wu, H. Shi, Longitudinal assessment of COVID-19 using a deep learning–based quantitative CT pipeline: illustration of two cases, Radiol.: Cardiothorac. Imaging 2 (2) (2020) e200082.

[8] Y. Bouchareb, P.M. Khaniabadi, F. Al Kindi, H. Al Dhuhli, I. Shiri, H. Zaidi, A. Rahmim, Artificial intelligence-driven assessment of radiological images for COVID-19, Comput. Biol. Med. 136 (2021) 104665.

[9] M. Sharafudeen, et al., Detecting skin lesions fusing handcrafted features in image network ensembles, Multimedia Tools Appl. (2022) 1–21.

[10] G. Han, X. Liu, H. Zhang, G. Zheng, N.Q. Soomro, M. Wang, W. Liu, Hybrid resampling and multi-feature fusion for automatic recognition of cavity imaging sign in lung CT, Future Gener. Comput. Syst. 99 (2019) 558–570.

[11] Y. Jin, D. Han, H. Ko, Memory-based semantic segmentation for off-road unstructured natural environments, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2021, pp. 24–31.

[12] R. Fusco, R. Grassi, V. Granata, S.V. Setola, F. Grassi, D. Cozzi, B. Pecori, F. Izzo, A. Petrillo, Artificial intelligence and COVID-19 using chest CT scan and chest X-ray images: machine learning and deep learning approaches for diagnosis and treatment, J. Pers. Med. 11 (10) (2021) 993.

[13] K. Gao, J. Su, Z. Jiang, L.-L. Zeng, Z. Feng, H. Shen, P. Rong, X. Xu, J. Qin, Y. Yang, et al., Dual-branch combination network (DCN): Towards accurate diagnosis and lesion segmentation of COVID-19 using CT images, Med. Image Anal. 67 (2021) 101836.

[14] I. Laradji, P. Rodriguez, O. Manas, K. Lensink, M. Law, L. Kurzman, W. Parker, D. Vazquez, D. Nowrouzezahrai, A weakly supervised consistency-based learning method for covid-19 segmentation in ct images, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021, pp. 2453–2462.

[15] N. Paluru, A. Dayal, H.B. Jenssen, T. Sakinis, L.R. Cenkeramaddi, J. Prakash, P.K. Yalavarthy, Anam-Net: Anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images, IEEE Trans. Neural Netw. Learn. Syst. 32 (3) (2021) 932–946.

[16] J. Adler-Milstein, A.K. Jha, Sharing clinical data electronically: a critical challenge for fixing the health care system, JAMA 307 (16) (2012) 1695–1696.

[17] P. Sharma, F.E. Shamout, D.A. Clifton, Preserving patient privacy while training a predictive model of in-hospital mortality, 2019, arXiv preprint arXiv:1912.00354.

[18] Y. Jiang, H. Chen, M. Loew, H. Ko, COVID-19 CT image synthesis with a conditional generative adversarial network, IEEE J. Biomed. Health Inf. 25 (2) (2020) 441–452.

[19] Y. Xu, A. Hosny, R. Zeleznik, C. Parmar, T. Coroller, I. Franco, R.H. Mak, H.J. Aerts, Deep learning predicts lung cancer treatment response from serial medical imaging, Clin. Cancer Res. 25 (11) (2019) 3266–3275.

[20] C. Jin, W. Chen, Y. Cao, Z. Xu, Z. Tan, X. Zhang, L. Deng, C. Zheng, J. Zhou, H. Shi, et al., Development and evaluation of an artificial intelligence system for COVID-19 diagnosis, Nature Commun. 11 (1) (2020) 1–14.

[21] S.G. Armato III, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, B. Zhao, D.R. Aberle, C.I. Henschke, E.A. Hoffman, et al., The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans, Med. Phys. 38 (2) (2011) 915–931.

[22] Y. Wang, Y. Zhang, Y. Liu, J. Tian, C. Zhong, Z. Shi, Y. Zhang, Z. He, Does non-COVID-19 lung lesion help? investigating transferability in COVID-19 CT image segmentation, Comput. Methods Programs Biomed. 202 (2021) 106004.

[23] Y.-J. Zhang, W.-J. Yang, D. Liu, Y.-Q. Cao, Y.-Y. Zheng, Y.-C. Han, R.-S. Jin, Y. Han, X.-Y. Wang, A.-S. Pan, et al., COVID-19 and early-stage lung cancer both featuring ground-glass opacities: a propensity score-matched study, Transl. Lung Cancer Res. 9 (4) (2020) 1516.

[24] H. Guan, M. Liu, Domain adaptation for medical image analysis: a survey, 2021, arXiv preprint arXiv:2102.09508.

[25] H. Chen, Y. Jiang, M. Loew, H. Ko, Unsupervised domain adaptation based COVID-19 CT infection segmentation network, Appl. Intell. (2021) 1–14.

[26] Y. Pathak, P.K. Shukla, A. Tiwari, S. Stalin, S. Singh, Deep transfer learning based classification model for COVID-19 disease, Irbm (2020).

[27] U. Özkaya, Ş. Öztürk, M. Barstugan, Coronavirus (COVID-19) classification using deep features fusion and ranking technique, in: Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach, Springer, 2020, pp. 281–295.

[28] S.S. Verma, A. Prasad, A. Kumar, CovXmlc: High performance COVID-19 detection on X-ray images using Multi-Model classification, Biomed. Signal Process. Control 71 (2022) 103272.

[29] A. Bhattacharyya, D. Bhaik, S. Kumar, P. Thakur, R. Sharma, R.B. Pachori, A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images, Biomed. Signal Process. Control 71 (2022) 103182.

[30] X. Ouyang, J. Huo, L. Xia, F. Shan, J. Liu, Z. Mo, F. Yan, Z. Ding, Q. Yang, B. Song, et al., Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia, IEEE Trans. Med. Imaging 39 (8) (2020) 2595–2605.

[31] A. Oulefki, S. Agaian, T. Trongtirakul, A.K. Laouar, Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images, Pattern Recognit. 114 (2021) 107747.

[32] N.S. Punn, S. Agarwal, Chs-net: A deep learning approach for hierarchical segmentation of covid-19 via ct images, Neural Process. Lett. (2022) 1–22.

[33] T. Zhou, S. Canu, S. Ruan, An automatic covid-19 ct segmentation based on u-net with attention mechanism, 2020, arXiv preprint arXiv:2004.06673.

[34] R. Karthik, R. Menaka, M. Hariharan, D. Won, Contour-enhanced attention CNN for CT-based COVID-19 segmentation, Pattern Recognit. 125 (2022) 108538.

[35] H. Hu, L. Shen, Q. Guan, X. Li, Q. Zhou, S. Ruan, Deep co-supervision and attention fusion strategy for automatic COVID-19 lung infection segmentation on CT images, Pattern Recognit. 124 (2022) 108452.

[36] D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, L. Shao, Inf-net: Automatic covid-19 lung infection segmentation from ct images, IEEE Trans. Med. Imaging 39 (8) (2020) 2626–2637.

[37] X. Liu, Q. Yuan, Y. Gao, K. He, S. Wang, X. Tang, J. Tang, D. Shen, Weakly supervised segmentation of COVID19 infection with scribble annotation on CT images, Pattern Recognit. 122 (2022) 108341.

[38] A. Malhotra, S. Mittal, P. Majumdar, S. Chhabra, K. Thakral, M. Vatsa, R. Singh, S. Chaudhury, A. Pudrod, A. Agrawal, Multi-task driven explainable diagnosis of COVID-19 using chest X-ray images, Pattern Recognit. 122 (2022) 108243.

[39] Y. Jiang, H. Chen, H. Ko, D.K. Han, Few-shot learning for CT scan based COVID-19 diagnosis, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2021, pp. 1045–1049.

[40] S. Qiao, W. Shen, Z. Zhang, B. Wang, A. Yuille, Deep co-training for semi-supervised image recognition, in: Proceedings of the European Conference on Computer Vision (Eccv), 2018, pp. 135–152.

[41] Y. Yang, S. Soatto, Fda: Fourier domain adaptation for semantic segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 4085–4095.

[42] Q. Li, L. Liu, Artificial intelligence-based semisupervised self-training algorithm in pathological Tissue image segmentation, Comput. Intell. Neurosci. 2022 (2022).

[43] X. Luo, M. Hu, T. Song, G. Wang, S. Zhang, Semi-supervised medical image segmentation via cross teaching between CNN and transformer, 2021, arXiv preprint arXiv:2112.04894.

[44] X. Chen, Y. Yuan, G. Zeng, J. Wang, Semi-Supervised Semantic Segmentation with Cross Pseudo Supervision, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 2613–2622.

[45] Y. Ouali, C. Hudelot, M. Tami, Semi-supervised semantic segmentation with cross-consistency training, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 12674–12684.

[46] Y. Zhang, L. Yang, J. Chen, M. Fredericksen, D.P. Hughes, D.Z. Chen, Deep adversarial networks for biomedical image segmentation utilizing unannotated images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2017, pp. 408–416.

[47] T.-H. Vu, H. Jain, M. Bucher, M. Cord, P. Pérez, Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 2517–2526.

[48] V. Verma, K. Kawaguchi, A. Lamb, J. Kannala, A. Solin, Y. Bengio, D. Lopez-Paz, Interpolation consistency training for semi-supervised learning, Neural Netw. 145 (2022) 90–106.

[49] A. Tarvainen, H. Valpola, Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results, 2017, arXiv preprint arXiv:1703.01780.

[50] C. Bucilu, R. Caruana, A. Niculescu-Mizil, Model compression, in: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006, pp. 535–541.

[51] L. Yu, S. Wang, X. Li, C.-W. Fu, P.-A. Heng, Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 605–613.

[52] W. Hang, W. Feng, S. Liang, L. Yu, Q. Wang, K.-S. Choi, J. Qin, Local and global structure-aware entropy regularized mean teacher model for 3d left atrium segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2020, pp. 562–571.

[53] Y. Jin, D. Han, H. Ko, Trseg: Transformer for semantic segmentation, Pattern Recognit. Lett. 148 (2021) 29–35.

[54] Y. Xu, B. Du, L. Zhang, Q. Zhang, G. Wang, L. Zhang, Self-ensembling attention networks: Addressing domain shift for semantic segmentation, in: Proceedings of the AAAI Conference on Artificial Intelligence, 33 (2019) 5581–5588.

[55] J. Choi, T. Kim, C. Kim, Self-ensembling with gan-based data augmentation for domain adaptation in semantic segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 6830–6840.

[56] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[57] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, C. Pal, The importance of skip connections in biomedical image segmentation, in: Deep Learning and Data Labeling for Medical Applications, Springer, 2016, pp. 179–187.

[58] http://medicalsegmentation.com/covid19/.

[59] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., Pytorch: An imperative style, high-performance deep learning library, Adv. Neural Inf. Process. Syst. 32 (2019).

[60] C.S.Perone, cclauss, E.Saravia, P.L.Ballester, MohitTare, perone/medicaltorch: Release v0.2, 2018, 2018, https://zenodo.org/record/1495335#.YtQAK-xBxFM.

[61] M.L. Waskom, Seaborn: statistical data visualization, J. Open Source Softw. 6 (60) (2021) 3021, http://dx.doi.org/10.21105/joss.03021.

[62] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[63] J. Bertels, T. Eelbode, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, M.B. Blaschko, Optimizing the dice score and jaccard index for medical image segmentation: Theory and practice, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 92–100.

[64] C. Zhang, X. Shen, H. Cheng, Q. Qian, Brain tumor segmentation based on hybrid clustering and morphological operations, Int. J. Biomed. Imaging 2019 (2019).

[65] C. Kaushal, K. Kaushal, A. Singla, Firefly optimization-based segmentation technique to analyse medical images of breast cancer, Int. J. Comput. Math. 98 (7) (2021) 1293–1308.

[66] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (11) (2008).