



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Dynamic feature learning for COVID-19 segmentation and classification

Xiaoqin Zhang^{a,*}, Runhua Jiang^a, Pengcheng Huang^a, Tao Wang^a, Mingjun Hu^a,
Andrew F. Scarsbrook^{b,c}, Alejandro F. Frangi^{d,e}

^a College of Computer Science and Artificial Intelligence, Wenzhou University, China

^b Radiology Department, Leeds Teaching Hospitals NHS Trust, UK

^c Leeds Institute of Medical Research, University of Leeds, UK

^d Centre for Computational Imaging and Simulation Technologies in Biomedicine, Leeds Institute for Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK

^e Department of Electrical Engineering, Department of Cardiovascular Sciences, KU Leuven, Belgium

ARTICLE INFO

Keywords:

COVID-19

Computed tomography

Dynamical fusion

Transfer learning

ABSTRACT

Since December 2019, coronavirus SARS-CoV-2 (COVID-19) has rapidly developed into a global epidemic, with millions of patients affected worldwide. As part of the diagnostic pathway, computed tomography (CT) scans are used to help patient management. However, parenchymal imaging findings in COVID-19 are non-specific and can be seen in other diseases. In this work, we propose to first segment lesions from CT images, and further, classify COVID-19 patients from healthy persons and common pneumonia patients. In detail, a novel Dynamic Fusion Segmentation Network (DFSN) that automatically segments infection-related pixels is first proposed. Within this network, low-level features are aggregated to high-level ones to effectively capture context characteristics of infection regions, and high-level features are dynamically fused to model multi-scale semantic information of lesions. Based on DFSN, Dynamic Transfer-learning Classification Network (DTCN) is proposed to distinguish COVID-19 patients. Within DTCN, a pre-trained DFSN is transferred and used as the backbone to extract pixel-level information. Then the pixel-level information is dynamically selected and used to make a diagnosis. In this way, the pre-trained DFSN is utilized through transfer learning, and clinical significance of segmentation results is comprehensively considered. Thus DTCN becomes more sensitive to typical signs of COVID-19. Extensive experiments are conducted to demonstrate effectiveness of the proposed DFSN and DTCN frameworks. The corresponding results indicate that these two models achieve state-of-the-art performance in terms of segmentation and classification.

1. Introduction

The world is experiencing a global pandemic due to the outbreak of coronavirus 2019 (COVID-19). This outbreak is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As reported by the Center for Systems Science and Engineering (to data June 26, 2020), there have been 9,586,769 confirmed cases and 488,824 deaths worldwide. Reverse-transcription polymerase chain reaction (RT-PCR) is the gold standard for diagnosis of COVID-19. However, RT-PCR is time-consuming and labour-intensive. Therefore, the radiological imaging techniques, particularly chest radiography (CXR) and computed tomography (CT), have emerged as an important complement to RT-PCR.

Among the radiological imaging techniques, CT screening is widely

researched due to its low energy consumption and anatomical information of the lungs. In recent researches [1,2], it has been demonstrated that the parenchymal imaging findings of COVID-19 pneumonia on CT scans include multi-focal peripheral ground-glass opacity (GGO) and pulmonary consolidation, which respectively appear in the early and late stages of COVID-19. Therefore, automated evaluation of CT scans can be helpful in fighting against COVID-19.

Many deep learning methods have been proposed and achieved excellent performance [3]. For example, Fan et al. [4] use a parallel partial decoder to aggregate high-level features, and propose a semi-supervised network to segment COVID-19 CT scans. In Ref. [5], a COVID-Net is proposed to diagnose COVID-19 patients from CXR. Zhang et al. [6] propose an anomaly detection model to analyse CXRs in COVID-19. Most existing research only focuses on segmenting lesions or

* Corresponding author.

E-mail address: zhangxiaoqinnan@gmail.com (X. Zhang).

classifying CT images, and may be insufficient for assisting clinicians. This is because, clinicians first review CT images to identify any lesion and their location, and then make decisions of the likely diagnosis based on the location and morphology of lesions. This indicates that segmentation and classification tasks are closely interrelated in clinical practice, and the segmentation task can provide necessary information for classification. Therefore, we propose a COVID-19 diagnosis methodology that employs both segmentation and classification tasks. However, as discussed in previous works, COVID-19 segmentation in CT scans is still a challenging task due to two issues: 1) Multi-scale context information is not fully utilized [7], 2) The high variance in texture, size and position of infections are challenging for segmentation, and inter-class variance of lesions are small [4]. On the other hand, the already challenging overlap in appearances between the different causes of pneumonia are always further complicated by limited image resolution [8].

To address the above issues, this work proposes the Dynamic Fusion Segmentation Network (DFSN) and Dynamic Transfer-learning Classification Network (DTCN), and Fig. 1 provides a rough view of proposed networks. On the one hand, segmentation-related issues are addressed by two modules in DFSN, namely the inter-stage and intra-stage fusion module. To address the first issue of utilizing context information, several inter-stage fusion modules are employed to combine low-level features with high-level ones, thus providing global information guidance flow for generating segmentation results. In contrast to methods that use skip-connections to fuse features from symmetrical layers (*i.e.*, from front or end of networks, layers have the same order) [7,9], the inter-stage fusion module takes features from unsymmetrical layers as inputs. It adopts convolution layers with large kernels to estimate such guidance. Besides, high-level features are dynamically fused by an intra-stage module to address issues of lesion characteristics (*e.g.*, texture, size and position). This module first takes several pixel-wise convolution layers to estimate location-specific fusion weights, then uses these weights to fuse features related to different lesions. Based on DFSN, a transfer-learning-based network (*i.e.*, DTCN) is then proposed to address the issue of distinguishing signs of diseases. Concretely, DTCN first utilizes a DFSN with transferred weights to extract pixel-level information, then adaptively selects this information according to clinical knowledge, and finally, generates image labels. In this way, parenchymal changes due to infection can be comprehensively estimated, and by considering their clinical meanings, accurate classifications can be easily achieved.

The main contributions of this work can be summarized as follows:

- We propose a novel Dynamic Fusion Segmentation Network (DFSN) for COVID-19 segmentation. Inside this network, inter-stage and intra-stage fusion modules are employed to fuse multi-scale context information and semantic information. To the best of our knowledge, this is the first work to consider differences of low-level and high-level features for designing fusion methodologies.
- Based on DFSN, we further introduce a novel Dynamic Transfer-learning Classification Network (DTCN) for classifying COVID-19 patients. Inside DTCN, a pre-trained DFSN is transferred as the backbone to extract pixel-level semantic information. Through systematic experiments, we also demonstrate the clinical significance of pixel-level information and their importance for COVID-19 diagnosis. According to our best knowledge, this is the first work to consider clinical significance of segmentation results for designing deep networks.
- We evaluate the proposed methods through comprehensive experiments. The results demonstrate that our methods achieve state-of-the-art performance in terms of segmentation and classification.

2. Related work

In this section, we briefly review several researches closely related to this work, including segmentation in CT, classification in CT and deep learning for COVID-19.

2.1. Segmentation in CT

Compared with RT-PCR, CT imaging is a more popular technique for the diagnosis and assessment of lung diseases [10,11]. As observed in Refs. [12,13], by segmenting organs and lesions from CT scans, doctors can quickly obtain information crucial for diagnosing lung diseases. Following this observation, many segmentation methods have been proposed and achieved significant performance. For example, in Ref. [14], the support vector machine (SVM) classifier is proposed to segment lung nodules from CT images. Shen et al. [15] present an automated lung segmentation system by taking the bidirectional chain code. Though this system achieves better performance than its counterparts, it still suffers from similar visual appearances of nodules and other tissues. To address this issue, methods based on convolutional

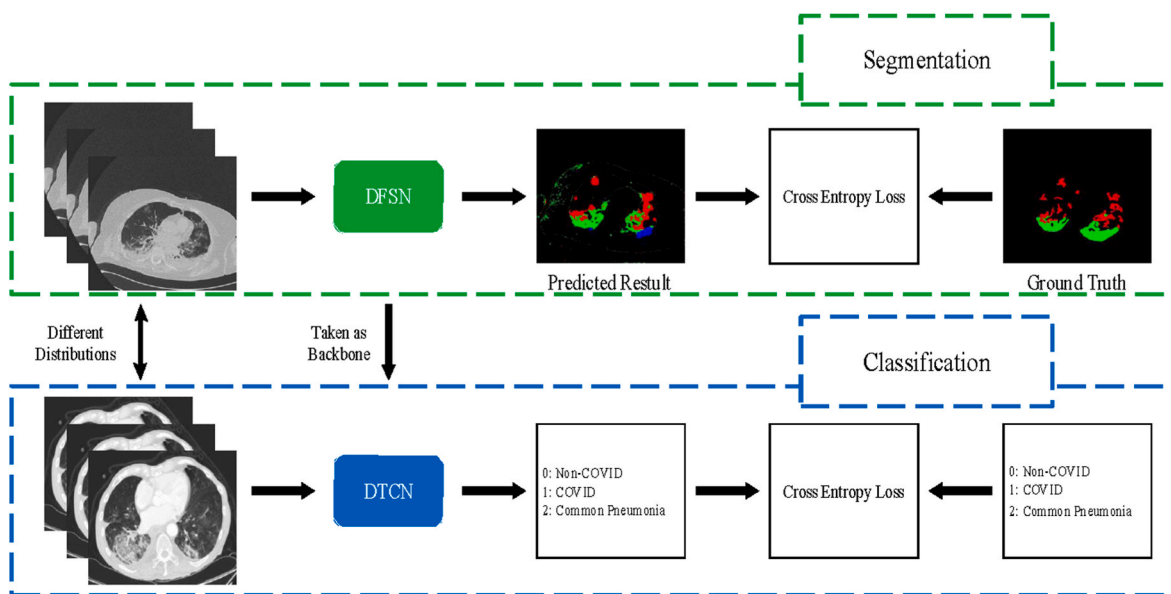


Fig. 1. The proposed DFSN is trained to generate semantic segmentation results accurately. In the proposed DTCN, DFSN is transferred as the backbone and fine-tuned in datasets with different distributions.

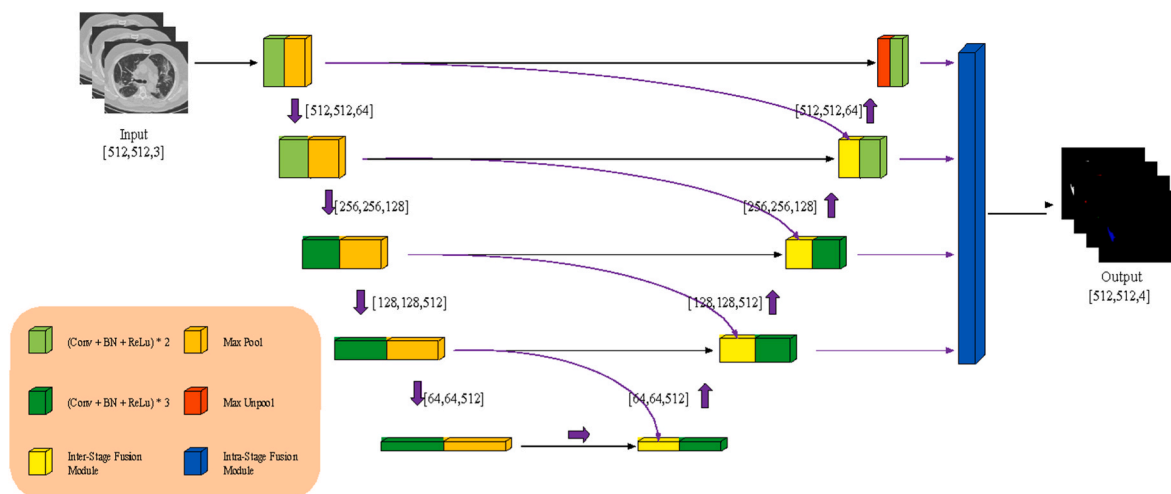


Fig. 2. The overall architecture of the proposed DFSN.

neural networks are widely studied. Wang et al. [16] introduce a central focused convolutional neural network to extract nodule-sensitive features from both 2D and 3D CT images. Jin et al. [17] utilize generative adversarial network (GAN) with multi-mask reconstruction loss to improve the robustness of progressive holistically nested network (P-HNN). According to their experiments, additional data, synthesized by GAN, is beneficial to overall segmentation performance. Jiang et al. [18] add multiple residual streams of varying resolutions and propose two networks to segment lung tumours. For further capturing global and multi-scale context information, Feng et al. [7] combine two pyramid convolutional modules and propose the Context Pyramid Fusion Network (CPFNet). In this work, we utilize the above context information further by employing two different feature fusion modules. Additionally, via comprehensive experiments, we demonstrate that, for an encoder-decoder network such as UNet and CPFNet, fusing features from the encoder is beneficial to locate infected regions, and fusing features from the decoder is beneficial to estimate pixel labels.

2.2. Classification in CT

Apart from segmenting organs and lesions on CT scans, computer-aided diagnosis (CAD) systems are also thought to be useful in effectively classifying multi-category CT images. For example, Li et al. [19] design a 2D convolutional neural network (CNN) to distinguish CT images in three categories, *i.e.*, COVID-19 pneumonia, community acquired pneumonia (CAP) and non-pneumonia. Specifically, input CT slices are fed through a pre-trained ResNet50 [20] to extract features. Then these features are combined and fed through a fully connected layer. In Ref. [8], Kang et al. conduct the classification task with multi-view presentation learning, which are achieved by encoding information from different aspects of features. In contrast, Wang et al. [21] combine two 3D-ResNets [20] and use a prior-attention strategy to guide them to learn more discriminative representations for the pneumonia-type classification. In Ref. [22], a 3D segmentation model is first used to segment lesion locations. Then, all locations are separately classified, and the overall classification result is obtained by using the Noisy-or-Bayesian function. In this work, we conduct a three-category classification task by transfer learning, which is achieved by fine-tuning a pre-trained segmentation model (*i.e.*, DFSN). In addition, through systematic experiments, we observe the clinical significance of segmentation results is vital for classifying on COVID-19 CT scans.

2.3. Deep learning for COVID-19

Deep learning methods have been widely employed in many CAD

systems for COVID-19 [3]. Joseph et al. [23] categorize these methods into three classes: patient scale (*e.g.*, medical imaging for diagnosis), molecular scale (*e.g.*, protein structure prediction) and population-scale (*e.g.*, epidemiology). In this Section, we briefly discuss patient-specific methods. In Ref. [24], a Spatial Transformer Network is proposed to predict disease severity in lung ultrasonography and localize pathological signatures. In Ref. [25], FC-DenseNet is first used to provide segmentation results. Based on this result, different patches of a CXR are cropped and fed into different classification networks. Then, the final classification result is obtained by majority voting from the results of cropped patches. Wu et al. [26] propose a Joint Classification and Segmentation (JCS) system to perform classification and segmentation of COVID-19 CT images. However, the JCS system performs the segmentation diagnosis only if the classification results are COVID-19 predictions. In Ref. [27], the authors proposed an AI system to conduct both classification, segmentation, and quantitative measurement tasks. Within this system, input images are initially put through a segmentation network to obtain lung-lesion maps. Then, the maps are taken as the input of a classification network for generating image labels, which are finally incorporated with clinical metadata to make quantitative measurements. Similar to Ref. [26], authors of [28] proposed a multi-task model, which contains parallel classification and segmentation branches and is supervised by a multi-task loss function.

Besides, several network architectures and mechanisms are considered in developing COVID-19 diagnosis models. For example, given an input image, previous [9] takes an encoder and a recurrent decoder to provide pixel-level diagnosis, while ResNet [20] and DenseNet [29] both utilize residual blocks and skip connections to make image-level classification. However, because high-level features tend to lose details of the input image, the above methods are easy prone to failure on complicated imaging data. For solving this issue, mechanisms such as attention mechanisms [30], multi-view presentation learning [8] and semi-supervision [4] are employed into the above models. By contrast, in this work, we first explore characteristics of multi-level features, and propose separate fusion blocks for features in different levels. As a result, the proposed DFSN can segment CT images more accurately. In addition, DFSN is used to form the transfer learning, which aims at diagnosing CT images. During this process, medical knowledge of different lesions is utilized to form model architectures.

3. Proposed method

In this Section, we first introduce the proposed Dynamic Fusion Segmentation Network (DFSN). Then, we discuss the motivation of integrating transfer learning and feature selection within a joint

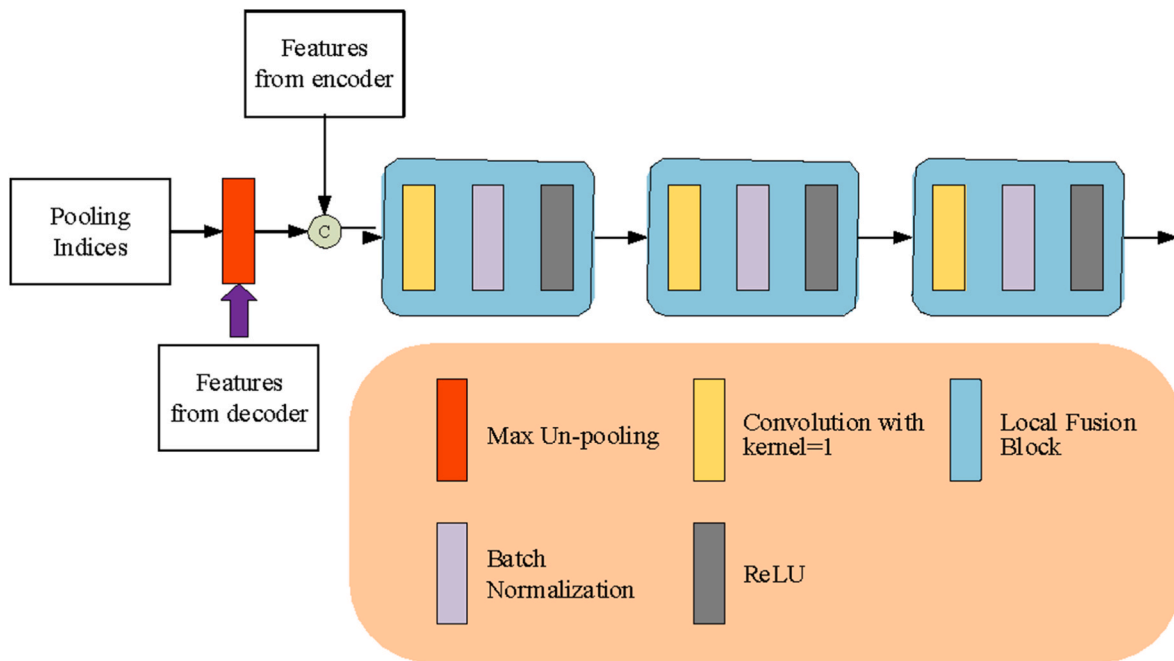


Fig. 3. The architecture of the proposed inter-stage fusion module has three inputs. It takes several pixel-wise convolution layers to fuse these inputs.

framework, which we coin Dynamic Transfer-learning Classification Network (DTCN).

3.1. Dynamic Fusion Segmentation Network

The architecture of the proposed DFSN is shown in Fig. 2. A CT image is first stacked on a three-dimensional tensor for alleviating the rapid increase of channels. After that, the three-dimensional tensor is fed through a U-shape model to extract features with different channels and resolutions. The tensor is first fed to an encoder (the five blocks on the left) to extract context information. Then encoder-generated features are fed to a decoder (the five blocks on the right) to obtain semantic information. Inside the encoder and decoder, feature maps are processed by several max pooling and convolution layers. Despite their demonstrated ability to estimate context information, this architecture blurs fine detail such as lesion boundaries, leading to poor segmentation performance [7]. Inspired by existing methods [4], features at different levels are fused to avoid this problem. In addition, by considering characteristics of multi-level features [31], two feature fusion modules are designed. For convenience, they are named as inter-stage and intra-stage fusion modules, and will be detailed in following subsections.

3.1.1. Inter-stage fusion module

As discussed in Ref. [31], the encoder can learn context information, including boundaries and category characteristics of objects. However, this information may be progressively weakened after going through deep layers. Besides, the simple skip-connection between the encoder and decoder is an indiscriminate combination, which always introduces irrelevant clutter [7]. Therefore, in this work, we propose the inter-stage fusion module to better use context information.

Fig. 3 shows the proposed inter-stage fusion module, which has three different inputs, namely pooling indices from max-pooling layers, features from the encoder and the decoder. For convenience, these three inputs are denoted as P_b , F_e and F_d , respectively. Inside this module, the P_b is used to magnify resolutions of F_d , which contains more semantic information than F_e [32]. After that, F_d is concatenated with F_e , thus integrating the context information extracted by the encoder and the semantic information estimated by the decoder. The three local fusion blocks comprising pixel-wise convolution, batch-normalization, and

ReLU layers, provide semantic and contextual information into the decoder. The proposed module differs from Refs. [4,9,29] for the following reasons: 1) The fused features are different. To be specific, Ronneberger et al. [9] fuse the encoder and decoder features in a balanced manner (e.g., in Fig. 2, features generated by the left-upper block are combined to the right-upper block), while the proposed module fuses features transported by the purple arrows. 2) Their fusion methods are different. The method proposed in Ref. [9] directly adds different features, and the model proposed by Ref. [7] fuses features through diluted convolution layers with 3×3 kernel. In contrast, the proposed module only takes pixel-wise convolution layers, with a 1×1 kernel.¹ Comprehensive experiments are conducted and presented in Section 4.3.1, which demonstrate the effectiveness of the intra-stage fusion module. According to the results, it is demonstrated that pixel-wise convolution layers can effectively utilize context information.

3.1.2. Intra-stage fusion module

As we mentioned in the introduction, infection characteristics such as texture, size and position are highly variable. At the same time, the inter-class variance of lesions is small. However, semantic edge information in high-level features help locate lesions and can provide useful constraints to guide label estimation for segmentation [4,33–35]. Thus, in this subsection, we propose an intra-stage fusion module to generate segmentation results dynamically. Four feature maps from the decoder are fused by the proposed intra-stage fusion module, whose architecture is shown in Fig. 4. As the resolutions and channels of these inputs are different from each other, they are first processed by feature normalization blocks with different parameters. The feature normalization blocks contain a pixel-wise convolution layer, a batch-normalization layer and a transposed convolution layer. The former two layers aim at reducing feature channels and covariate shift [36], while the last one recovers feature resolutions. Then, for generating location-adaptive fusion weights, output features of the Feature Normalization-7 are further fed into the Adaptive Weight Learner proposed in Ref. [32]. Specifically, this learner is formed by three pairs of convolution, batch

¹ For simplicity, batch normalization and ReLU layers of [7] and the proposed module are ignored.

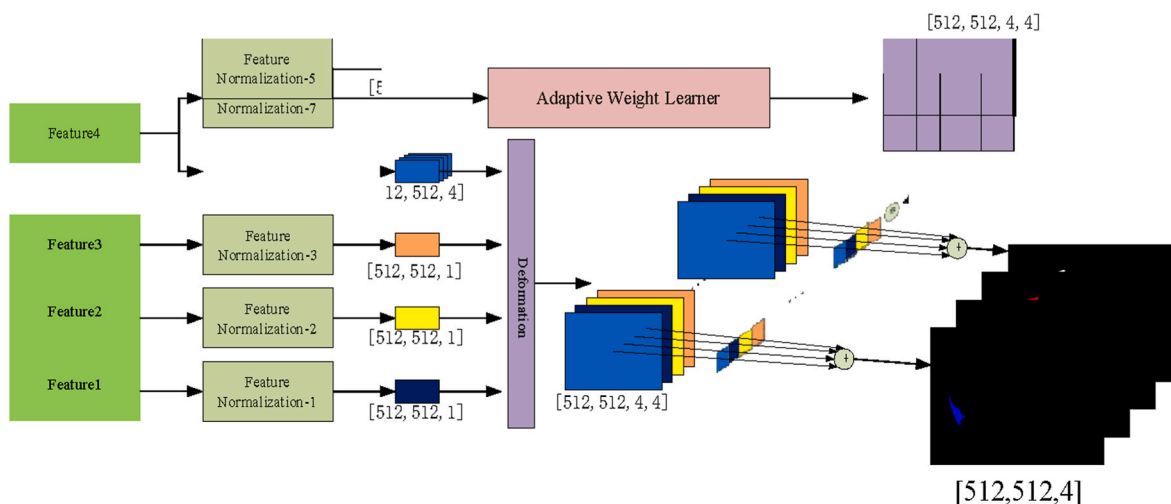


Fig. 4. The architecture of the proposed intra-stage fusion module. Multi-level features from the decoder are taken as input and processed by different feature normalization blocks. After that, the output features of these block are dynamically fused.

normalization and ReLU layers. A notable difference is that, in this work, the kernel size of these convolution layers is set to 5×5 . This is because lesion contours are blurrier than boundaries in healthy tissues [7], and compared with pixel-wise convolution, convolution layers with big receptive fields (e.g., big kernels) are more robust to the blurry contours [37]. After that, output features of Adaptive Weight Learner and the bottom four feature normalization layers are deformed to two 4D tensors by different methods. Specifically, output features of the Adaptive Weight Learner are deformed according to channel numbers. In contrast, features of the Feature Normalization-5 are first concatenated with other features, then, these four concatenated features are further concatenated to obtain the 4D tensors (indicated by coloured rectangles in Fig. 4). Finally, these two 4D tensors are multiplied to generate segmentation results (i.e., the right-bottom tensor). Overall, the intra-stage fusion module dynamically fuses features by the Adaptive Weight Learner, whose kernel sizes are different from Ref. [32]. To demonstrate the influence of this different configuration, several experiments are conducted in Section 4.3.1.

3.2. Dynamic Transfer-learning classification network

Fig. 5 provides an overview of the proposed Dynamic Transfer-learning Classification Network (DTCN), which takes a pre-trained semantic segmentation network to conduct the image classification task. To be specific, inside DTCN, a pre-trained DFSN is taken as a backbone to extract pixel-level information (i.e., segmentation results presented by features). Then, these extracted features are adaptively selected according to medical knowledge. Finally, an adaptive pooling layer reduces feature resolutions, and a fully connecting layer generates categories of the input image. The motivation to use DTCN comes from considering that, during diagnosis, clinicians first scout CT images, then locate lesions regions, and finally, make decisions according to regional appearances and contours of the detected regions. Thus, the goal of the first step is similar to the semantic segmentation task that predicts pixel-level categories, and the other steps aim at making decisions. In addition, it is easy to find that the above methodology can also be utilized for analyzing other medical images, which indicates that the proposed DTCN is efficient and heuristic for other tasks. Therefore, semantic segmentation methods can provide classification models with pixel-level information [22]. However, whether the semantic segmentation methods should be pre-trained and whether segmentation results are beneficial to classification remains unclear. In the following subsections, we discuss the above two questions by detailing the proposed DTCN.

3.2.1. Joint segmentation and classification by transfer learning

It is widely known that semantic segmentation aims at estimating categories of each pixel, and image classification aims at obtaining image-level categories. Therefore, these two tasks are closely related to one another, and their complementary relationship has been explored in medical image analysis. To be specific, Wu et al. [26] considered the results of semantic segmentation and classification tasks are jointly considered during diagnosis. Zhang et al. [27] first utilize a segmentation network to generate lung-lesion features, then provide prognosis analysis by considering the lung-lesion features and clinical metadata. However, it can be noted that method proposed in Ref. [27] requires not only numerous manually segmented images, but also need multi-modal data such as clinical metadata. However, achieving these two requirements is difficult and time-consuming. In addition, as discussed in Refs. [38–40], fine-tuning pre-trained models in datasets with different distributions can improve the robustness and effectiveness of learned feature space. Therefore, a pre-trained segmentation network is needed to provide pixel-level information.

3.2.2. Feature selection

Fig. 5 shows that output features of DFSN form a 4D feature map. A simple method is to directly take this 4D feature map to make classifications. However, for a segmentation network (i.e., DFSN), values of each pixel represent probabilities of a specific class. According to channels with the highest probability, the category of a single pixel can be obtained. Therefore, it is easy to find that each channel contains information about a particular class, and not all channels are beneficial to diagnose CT images. For example, during the process of distinguishing COVID-19 CT scans, channels related to the background (e.g., tissues around the lung) are useless. In this work, DFSN is trained for segmenting pixels of four categories, including background, GGO, consolidation and pleural effusion. Therefore, in the 4D feature maps, the feature at the second channel relates to consolidation, which is characterized by a homogeneous increase in lung parenchymal attenuation [41]. However, as indicated in Ref. [42], the most frequently observed features of COVID-19 pneumonia are bilateral involvement, peripheral distribution with GGO, rather than consolidation. Besides, perhaps consolidations are small, and their appearances are similar to adjacent structures, resulting in false-negative detection [4]. Due to the above two reasons, segmentation of consolidation is likely to be less useful for classifying COVID-19 images. For demonstrating this observation, comprehensive experiments are conducted and presented in Section 4.3.2.

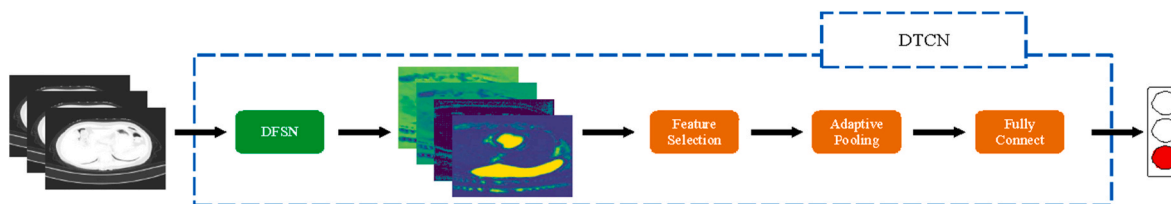


Fig. 5. The overall architecture of the proposed DTCN. The DTCN, with a CT scan as input, makes a classification using the transferred DFSN, feature selection layer, adaptive pooling layer, and the fully connected layer.

Table 1

Summary of the segmentation subsets.

Subset	Background	GGO	Consolidation	Pleural Effusion
Train	70	66	56	19
Test	30	30	22	6
Total	100	96	78	25

Table 2

Summary of the classification subsets.

Subset	Common Pneumonia	Non COVID-19	COVID-19
Train	200	234	191
Validation	100	58	60
Test	100	105	98
Total	400	387	349

4. Experiments

4.1. Data and metrics

4.1.1. Data for segmentation

As discussed in Ref. [4], only the COVID-19 CT Segmentation dataset, which consists of 100 axial CT images from different COVID-19 patients, is publicly accessible. In this dataset, all images are collected by the Italian Society of Medical and Interventional Radiology.² For identifying lung infections, all CT images are manually segmented by a radiologist, and all pixels are labeled by four categories, *i.e.*, background, GGO, consolidation and pleural effusion. For comparing DFSN and other methods, all images are randomly split into two subsets. Table 1 summarizes these two subsets. As discussed in Ref. [4], though this is the first open-access dataset for lung infection segmentation, it still suffers from the limited size and low-resolution images. Therefore, we also conduct a 5-fold experiments to further compare DFSN and other methods.

4.1.2. Data for classification

For comparing DTCN and other methods, we obtain a three-category dataset by combing COVID-CT-Dataset [43] and the dataset proposed in Ref. [27]. This approach ensures there are no duplicated images in the classification and segmentation datasets. In total, this synthetic dataset contains 1136 images, which are labeled as non-COVID-19, COVID-19 and common pneumonia. Images of common pneumonia are randomly extracted from the dataset described by Ref. [27], while images of other categories are taken from the COVID-CT-Dataset. As Table 2 presents, all images are split into train, validation and test subsets.

4.1.3. Metrics

In this work, the Intersection over Union (IoU), Dice coefficient (Dice), Precision (Pre) and Accuracy (Acc) are taken to evaluate performance of the proposed DFSN. For the three-category classification

task, the Accuracy (Acc), Precision (Pre) and Recall (Rec) are used to verify the classification ability of DTCN.

4.2. Implementation details

In this work, the proposed DFSN and DTCN are implemented on the PyTorch platform with NVIDIA Tesla K80 GPU. For training DFSN, the stochastic gradient descent (SGD) optimizer is used, in which momentum and learning rate are set to 0.9 and 0.001, respectively. During training, the batch size is fixed to 1, and the learning rate is decayed every 10 epochs. All training batches are randomly rotated for data augmentation. In addition, for alleviating the unbalanced segmentation problems, a weighted cross entry loss function is adopted, and these weights are 0.0013, 0.0261, 0.0501 and 1.000, respectively. On the other hand, the proposed DTCN is trained with the Adam optimizer [44] and an unweighted cross entry loss. In Adam, the first and second momentum values are fixed to 0.9 and 0.999, and the weight-decay is set to 0.1. The learning rate is set to 0.00001, and is dynamically decayed according to validation loss.

4.3. Model analysis

4.3.1. Analysis of segmentation

As discussed in Section 3.1, DFSN takes two modules to capture complementary relationships between multi-level features. Therefore, here we provide several baselines to demonstrate the effectiveness of the two fusion modules. Configurations of these baselines are shown in Table 3. Specially, as both SegNet [45] and the proposed DFSN take max pooling and un-pooling layers to change feature resolutions, we take SegNet as a baseline without the two proposed fusion modules, rather than the widely used UNet [9]. For the Addition-Fusion baseline, features from the encoder are additionally fused by the inter-stage fusion module.

Experimental results of these baselines are shown in Table 4. By comparing metrics of SegNet, Inter-Fusion, Intra-Fusion and Dual-Fusion, effectiveness of these two fusion modules is demonstrated. On the one hand, performance degradation of Addi-Fusion indicates that fusing multi-level features from the encoder is not beneficial for extracting context information. Similarly, the comparison result between Deeper-Fusion and Dual-Fusion indicates that taking deep modules to fuse multi-level features is ineffective. There are two reasons for this observation: 1) Deep fusion modules inevitably cause an increase in model parameters, which makes it hard to train with limited images; 2) Deep fusion modules can easily cause the gradient vanishing problem. On the other hand, features from different stages should be fused differently. That is, to capture multi-scale context information, pixel-wise convolutions are more effective than large-kernel convolutions. The latter, however, perform better at fusing semantic information.

4.3.2. Analysis of classification

Here we provide seven baselines to verify the performance of the transferred DFSN and feature selection layer. Table 5 provides a clear view of these seven baselines. Among them, maintained features of DFSN-Class₁₂₄₊ are additionally fused with the 2nd feature through a point-wise convolution layer. For the SegNet-Class, a pre-trained SegNet

² <https://www.sirm.org/category/senza-categoria/covid-19>.

Table 3

Configurations of segmentation baselines. For the Addition-Fusion baseline, features from encoder are additionally fused by the inter-stage fusion module.

Baselines	Inter-Stage Fusion	Inter-Stage Fusion	Depth	Kernels _{Inter}	Kernels _{Intra}
SegNet [45]	–	–	–	–	–
Inter-Fusion	✓	–	3	1-1-1	–
Intra-Fusion	–	✓	3	–	1-1-1
Dual-Fusion	✓	✓	3	1-1-1	1-1-1
Addi-Fusion	✓	✓*	3	1-1-1	1-1-1
Deeper-Fusion	✓	✓	6	1-1-1-1-1-1	1-1-1-1-1-1
Dual-Fusion-3	✓	✓	3	1-3-1	1-3-1
Dual-Fusion-5	✓	✓	3	1-5-1	1-5-1
Dual-Fusion-7	✓	✓	3	1-7-1	1-7-1
Dual-Fusion-5+	✓	✓	3	5-5-5	1-1-1
DFSN	✓	✓	3	1-1-1	5-5-5

Table 4

Results of segmentation baselines. The best and second-best results are respectively denoted by red and blue colors.

Baselines	IoU	Dice	Pre	Acc
SegNet	0.74	0.50	0.46	0.87
Inter-Fusion	0.78	0.49	0.46	0.87
Intra-Fusion	0.76	0.52	0.46	0.90
Dual-Fusion	0.79	0.51	0.47	0.90
Addi-Fusion	0.79	0.46	0.42	0.83
Deeper-Fusion	0.77	0.51	0.46	0.90
Dual-Fusion-3	0.77	0.47	0.44	0.84
Dual-Fusion-5	0.78	0.53	0.49	0.92
Dual-Fusion-7	0.77	0.49	0.46	0.84
Dual-Fusion-5+	0.79	0.46	0.43	0.83
DFSN	0.80	0.53	0.49	0.90

Table 5

Configurations of classification baselines.

Baselines	Transfer Learning	1st Feature	2nd Feature	3rd Feature	4th Feature
w/o transfer	–	✓	✓	–	✓
SegNet-Class	SegNet	✓	✓	–	✓
w/o selection	DFSN	✓	✓	✓	✓
DFSN-Class ₁₂₃	DFSN	✓	✓	✓	–
DFSN-Class ₁₃₄	DFSN	✓	–	✓	✓
DFSN-Class ₂₃₄	DFSN	–	✓	✓	✓
DFSN-Class ₁₂₄₊	DFSN	✓	✓	–	✓
DTCN	DFSN	✓	✓	–	✓

Table 6

Results of classification baselines. In the Pre and Rec columns, values in the three sub-columns are metrics related to certain categories (from left to right, non-COVID-19, COVID-19 and common pneumonia). The best and second best results are respectively denoted by red and blue colors.

Baselines	Acc	Pre	Rec				
w/o transfer	0.64	0.55	0.00	0.75	0.90	0.00	0.98
SegNet-Class	0.40	0.63	0.31	0.00	0.49	0.70	0.00
w/o selection	0.74	0.70	0.66	0.82	0.69	0.53	0.99
DFSN-Class ₁₂₃	0.65	0.56	0.00	0.78	0.93	0.00	1.00
DFSN-Class ₁₃₄	0.47	0.37	0.71	0.00	0.78	0.60	0.00
DFSN-Class ₂₃₄	0.56	0.00	0.67	0.51	0.00	0.72	1.00
DFSN-Class ₁₂₄₊	0.70	0.62	0.75	0.78	0.88	0.21	0.98
DTCN	0.77	0.73	0.76	0.82	0.76	0.57	0.98

Table 7

Results of compared segmentation methods. The best and second-best results are respectively denoted by red and blue colors.

Methods	IoU	Dice	Pre	Acc
SegNet [45]	0.74	0.50	0.46	0.87
UNet [9]	0.47	0.52	0.61	0.96
DeepLab v2 [46]	0.59	0.42	0.39	0.85
DeepLab v3 [47]	0.68	0.40	0.37	0.76
Attention UNet [30]	0.75	0.49	0.43	0.89
R2UNet [48]	0.31	0.24	0.25	0.83
Attention R2UNet ^a	0.47	0.34	0.33	0.79
InfNet-ResNet [4]	0.72	0.45	0.40	0.87
InfNet-VggNet [4]	0.76	0.38	0.35	0.87
Triage [28]	0.75	0.47	0.44	0.85
CopleNet [49]	0.74	0.50	0.48	0.88
DFSN	0.80	0.53	0.49	0.90

^a https://github.com/FENGShuanglang/Image_Segmentation.

is taken as the backbone.

Table 6 presents results of DTCN and the proposed baselines. By comparing results of w/o transfer and DTCN, it can be found that the pre-trained DFSN remarkably improves the overall performance. Results of SegNet-Class and DTCN also demonstrate the effectiveness of the transferred DFSN. The five baselines related to the feature selection layer indicate that features related to the consolidation are unhelpful to classification. In detail, comparing configurations of w/o selection, DFSN-Class₁₂₃, DFSN-Class₁₃₄, DFSN-Class₂₃₄ and DTCN, it can be found GGO are the most important features, and features related to the pleural effusion are secondarily important. In addition, by comparing DFSN-Class₁₂₄₊ and DTCN, effectiveness of the proposed feature selection layer is further demonstrated. Overall, the above comparison results demonstrate that: 1) Using a pre-trained segmentation network as the backbone of classification models is a promising method, 2) Segmentation results should be considered according to their clinical means.

4.4. Comparisons with other methods

4.4.1. Comparison of segmentation

In this subsection, the proposed DFSN is compared with several segmentation networks. As Table 7 presents, the compared methods include classical networks, such as SegNet [45], DeepLab v2 [46] and DeepLab v3 [47]. Furthermore, five segmentation networks designed for medical images and their variants are also compared. Inside UNet [9], an encoder-decoder architecture is taken to learn semantic information, and skip-connections are taken to fuse features from the encoder and decoder. Based on UNet, the Attention UNet [30] and R2UNet [48] are obtained by employing attention mechanisms and residual blocks. By combining these two models, the Attention R2UNet,³ which has shown

³ https://github.com/FENGShuanglang/Image_Segmentation.

Table 8

The 5-fold results of compared segmentation methods. The best and second-best results are respectively denoted by red and blue colors. Each metric is shown together with the standard deviation.

Methods	IoU	Dice	Pre	Acc
SegNet [45]	0.67 ± 0.13	0.47 ± 0.08	0.43 ± 0.07	0.86 ± 0.04
UNet [9]	0.71 ± 0.07	0.51 ± 0.03	0.50 ± 0.06	0.90 ± 0.05
DeepLab v2 [46]	0.60 ± 0.08	0.37 ± 0.10	0.36 ± 0.06	0.72 ± 0.16
DeepLab v3 [47]	0.68 ± 0.06	0.50 ± 0.04	0.45 ± 0.04	0.88 ± 0.04
Attention UNet [30]	0.75 ± 0.03	0.51 ± 0.06	0.47 ± 0.06	0.88 ± 0.04
R2UNet [48]	0.47 ± 0.11	0.32 ± 0.06	0.36 ± 0.07	0.80 ± 0.08
InfNet-ResNet [4]	0.68 ± 0.09	0.47 ± 0.06	0.42 ± 0.05	0.86 ± 0.03
InfNet-VggNet [4]	0.70 ± 0.03	0.51 ± 0.06	0.48 ± 0.07	0.88 ± 0.03
Triage [28]	0.63 ± 0.06	0.50 ± 0.04	0.50 ± 0.04	0.91 ± 0.02
JCS [26]	0.41 ± 0.06	0.35 ± 0.08	0.36 ± 0.08	0.88 ± 0.04
CopleNet [49]	0.64 ± 0.06	0.47 ± 0.05	0.46 ± 0.05	0.86 ± 0.08
DFSN	0.73 ± 0.01	0.52 ± 0.03	0.48 ± 0.02	0.88 ± 0.02

better segmentation performance, is obtained. For InfNet-ResNet and InfNet-VggNet, they are variants of InfNet [4]. As most methods in our comparison are trained with different datasets, all compared methods and DFSN are trained under the same protocol for a fair comparison.

Comparison results are shown in Table 7, Table 8 and Fig. 6. As can be seen from Table 7, the proposed DFSN achieves the best performance in terms of IoU and Dice. For other metrics (*i.e.*, Pre and Acc), DFSN also achieves state-of-the-art performance. In detail, compared with methods based on UNet, such as Attention UNet, R2UNet and Attention R2UNet, DFSN outperforms them by about 6%. This demonstrates that the proposed fusion modules are more effective than attention mechanisms used in Refs. [30,48]. Further, by comparing SegNet and DFSN, it can be found that context information is beneficial to segmentation, and effectively utilizing such information can remarkably improve segmentation performance. As Table 8 shows, during the 5-fold experiment, the proposed DFSN has the minimum standard deviation, which indicates that DFSN are more robust than these compare methods. On the

other hand, compared with recent methods such as Triage [28], JCS [26], and CopleNet [49], DFSN not only outperforms them in terms of values of most metrics, but also the standard deviation. The above comparison results demonstrate that DFSN achieves comparable performance and better stability than recent works. Fig. 6 further demonstrates the effectiveness of DFSN. It can be found that, compared with other methods, the results of DFSN are more similar to the ground truth. However, this figure also indicates that all methods wrongly classify background into consolidation or pleural effusion. There are two reasons for this phenomenon: 1) The adopted segmentation dataset is highly unbalanced (see Table 1), 2) Most of the CT images contain not only a large number of dark regions (*e.g.*, the corners of the bottom image), but also a broad and complex variation of tissues (*e.g.*, corners of the upper image).

Table 9

Results of compared classification methods. In the Pre and Rec columns, values in the three sub-columns are metrics related to certain categories (from left to right, non-COVID-19, COVID-19 and common pneumonia). The best and second-best results are respectively denoted by red and blue colors.

Methods	Acc	Pre			Rec		
EfficientNet [50]	0.65	0.71	0.51	0.80	0.19	0.79	0.99
DenseNet [29]	0.72	0.64	0.65	0.82	0.65	0.52	0.98
Inception-v3 [51]	0.37	0.41	0.34	0.40	0.27	0.53	0.33
ResNext-50 [52]	0.76	0.72	0.73	0.81	0.74	0.57	0.96
DarkCovidNet [53]	0.67	0.59	0.57	0.79	0.66	0.36	1.00
Jin et al. [54]	0.72	0.66	0.63	0.83	0.72	0.43	1.00
AdderNet [55]	0.52	0.52	0.45	0.58	0.43	0.40	0.75
GhostNet [56]	0.59	0.49	0.47	0.77	0.16	0.68	0.96
Res2Net [57]	0.65	0.58	0.51	0.81	0.56	0.42	0.98
Res2Next [57]	0.70	0.63	0.64	0.79	0.63	0.49	0.97
COVNet [19]	0.76	0.73	0.69	0.82	0.71	0.55	1.00
DTCN	0.77	0.73	0.76	0.82	0.76	0.57	0.98

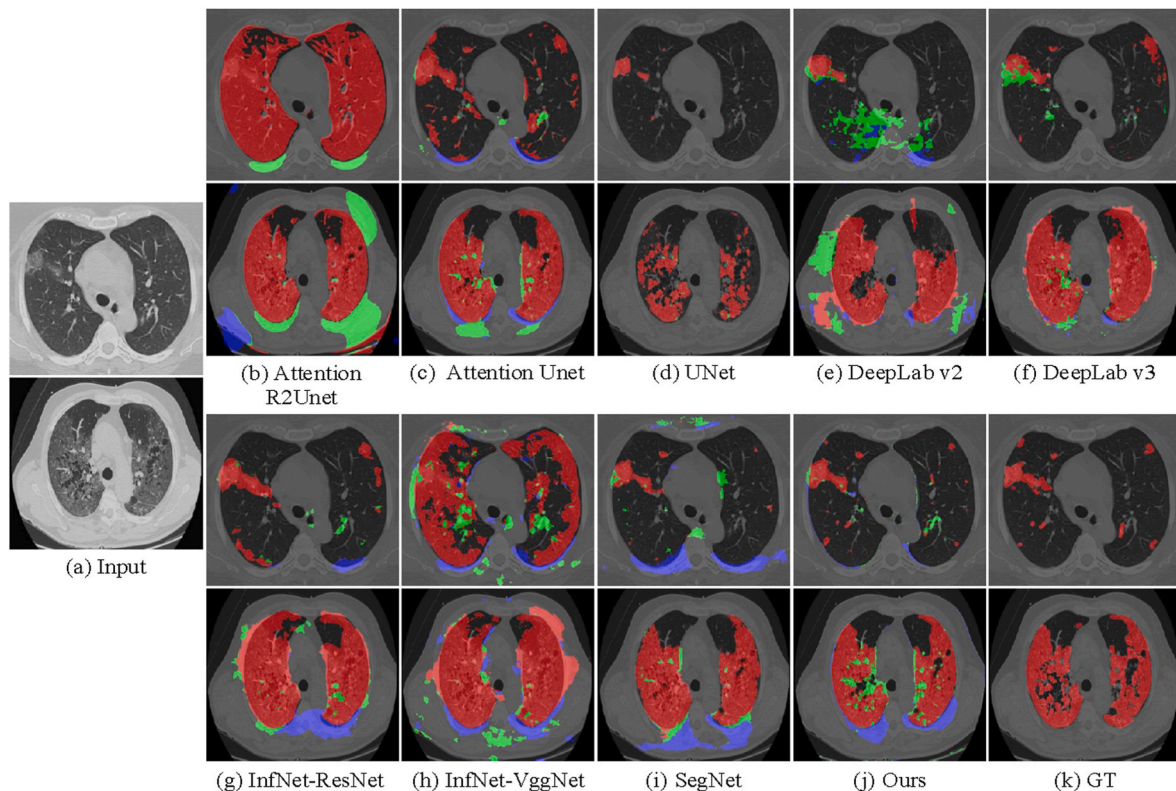


Fig. 6. Qualitative results of compared segmentation methods and the proposed DFSN. The red, green and blue pixels indicate GGO, consolidation and pleural effusion, respectively. **Best viewed in color.**

4.4.2. Comparison of classification

For demonstrating the performance of the proposed DTCN, six classification models are compared. Table 9 provides quantitative results of the six compared models and DTCN. Among these compared models, COVNet [19] and DarkCovidNet [53] are recently published methods for classifying COVID-19 images. For DenseNet [29] and ResNext-50 [52], the classification result is generated by using stacked convolution layers and skip-connections, which are widely used in both medical image analysis and traditional image classification. Other methods, e.g., EfficientNet [50], rely on efficient network design, such as multiple branches and convolution layers with different kernel sizes.

As Table 9 indicates, the proposed DTCN achieves the best performance in terms of Acc, Pre and Rec. Concretely, DTCN outperforms COVNet and DarkCovidNet by 0.01 and 0.1 in term of Acc. For classifying images of the first and second categories, the performance of DTCN indicates that pixel-level information aids the classification of Non-COVID-19 and COVID-19 CT scans. Thus DTCN can effectively locate lesions (e.g., GGO and pleural effusion) and make more accurate classifications than its counterparts. Besides, during the detection of COVID-19 CT scans, the Rec of DTCN is worse than EfficientNet. This indicates that DTCN makes more false-negative classifications than EfficientNet. By analyzing classification baselines of DTCN, DTCN and EfficientNet, the reasons can be attributed to that segmentation results related to the background frequently introduce irrelevant information. In addition, by comparing DTCN with recent models, it can be found that DTCN outperforms most of them. Specifically, compared with novel models such as AdderNet [55], GhostNet [56], and Res2Net [57], it can be seen that owing to the transfer learning and the feature selection, DTCN can achieve better performance with limited parameters and simple architectures. It can also be found that DTCN achieves comparable performance of recent works proposed for medical image analysis, e.g., Jin *et al.* [54] and COVNet [19]. Compared with [54], the proposed DTCN is mildly worse in term of Rec. However, by observing the middle column of Pre and Rec, it is demonstrated that DTCN can distinguish COVID-19 images more accurately than [54]. For the other work COVNet [19], DTCN also achieves better performance in terms of Acc and Pre. On the Rec metric, though DTCN does not outperform COVNet in classifying images of common pneumonia, it still achieves better performance in classifying images of non-COVID-19 and COVID-19.

5. Conclusion

In this work, we propose a novel Dynamic Fusion Segmentation Network, which adopts two fusion modules to improve the identification of lesion regions. Moreover, we also design a transfer-learning-based network (i.e., Dynamic Transfer-learning Classification Network) to distinguish COVID-19 CT images from Non-COVID and common pneumonia scans. This network employs a pre-trained segmentation network to extract pixel-level information (i.e., segmentation results), which makes classifications according to selected pixel-level information. For demonstrating effectiveness of DFSN and DTCN, extensive experiments are conducted on benchmark segmentation dataset and a three-class classification dataset. Experimental results demonstrate that the proposed models achieve not only state-of-the-art performance, but also have great potential in the clinical assessment of patients with suspected COVID-19.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation

of China [grant nos. 61922064, U2033210, 62101387].

References

- [1] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, et al., Correlation of chest ct and rtpcr testing in coronavirus disease 2019 (covid-19) in China: a report of 1014 cases, *Radiology* (2020), 200642.
- [2] Z. Ye, Y. Zhang, Y. Wang, Z. Huang, B. Song, Chest ct manifestations of new coronavirus disease 2019 (covid-19): a pictorial review, *Eur. Radiol.* (2020) 1–9.
- [3] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, et al., Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19, *IEEE Reviews in Biomedical Engineering* (2020).
- [4] D. Fan, T. Zhou, G. Ji, Y. Zhou, G. Chen, H. Fu, et al., Inf-net: automatic covid-19 lung infection segmentation from ct scans, *arXiv preprint arXiv* (2020), 200414133.
- [5] L. Wang, A. Wong, Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images, *arXiv* 2020 (2003) arXiv.
- [6] J. Zhang, Y. Xie, Y. Li, C. Shen, Y. Xia, Covid-19 screening on chest x-ray images using deep learning based anomaly detection, *arXiv preprint arXiv* (2020), 200312338.
- [7] S. Feng, H. Zhao, F. Shi, X. Cheng, M. Wang, Y. Ma, et al., Cpfnet: context pyramid fusion network for medical image segmentation, *IEEE Trans. Med. Imag.* (2020).
- [8] H. Kang, L. Xia, F. Yan, Z. Wan, F. Shi, H. Yuan, et al., Diagnosis of Coronavirus Disease 2019 (Covid-19) with Structured Latent Multi-View Representation Learning, *IEEE Transactions on Medical Imaging*, 2020.
- [9] Zhang Xiaoqin, Jiang Runhua, Wang Tao, et al., Recursive Neural Network for Video Deblurring, in: *IEEE Transactions on Circuits and Systems for Video Technology*, IEEE, 2021, pp. 3025–3036.
- [10] I. Sluimer, A. Schilham, M. Prokop, B. Van Ginneken, Computer analysis of computed tomography scans of the lung: a survey, *IEEE Trans. Med. Imag.* 25 (4) (2006) 385–405.
- [11] B. Kamble, S.P. Sahu, R. Doriya, A review on lung and nodule segmentation techniques, in: *Advances in Data and Information Sciences*, Springer, 2020, pp. 555–565.
- [12] P.M. Gordaliza, A. Muñoz-Barrutia, M. Abella, M. Desco, S. Sharpe, J.J. Vaquero, Unsupervised ct lung image segmentation of a mycobacterium tuberculosis infection model, *Sci. Rep.* 8 (1) (2018) 1–10.
- [13] A. Munoz-Barrutia, M. Ceresa, X. Artaechevarria, L.M. Montuenga, C. Ortiz-de Solorzano, Quantification of lung damage in an elastase-induced mouse model of emphysema, *Int. J. Biomed. Imag.* 2012 (2012).
- [14] M. Keshani, Z. Azimifar, F. Tajeripour, R. Boostani, Lung nodule segmentation and recognition using svm classifier and active contour modeling: a complete intelligent system, *Comput. Biol. Med.* 43 (4) (2013) 287–300.
- [15] S. Shen, A.A. Bui, J. Cong, W. Hsu, An automated lung segmentation approach using bidirectional chain codes to improve nodule detection accuracy, *Comput. Biol. Med.* 57 (2015) 139–149.
- [16] S. Wang, M. Zhou, Z. Liu, Z. Liu, D. Gu, Y. Zang, et al., Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation, *Med. Image Anal.* 40 (2017) 172–183.
- [17] D. Jin, Z. Xu, Y. Tang, A.P. Harrison, D.J. Mollura, Ct-realistic lung nodule simulation from 3d conditional generative adversarial networks for robust lung segmentation, in: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2018, pp. 732–740.
- [18] J. Jiang, Y. Hu, C. Liu, D. Halpenny, M.D. Hellmann, J.O. Deasy, et al., Multiple resolution residually connected feature streams for automatic lung tumor segmentation from ct images, *IEEE Trans. Med. Imag.* 38 (1) (2018) 134–144.
- [19] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, et al., Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct, *Radiology* (2020), 200905.
- [20] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [21] J. Wang, Y. Bao, Y. Wen, H. Lu, H. Luo, Y. Xiang, et al., Prior-attention residual learning for more discriminative covid-19 screening in ct images, *IEEE Trans. Med. Imag.* (2020).
- [22] C. Butt, J. Gill, D. Chun, B.A. Babu, Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia, *Applied Intelligence*, 2020, p. 1.
- [23] V. Rajinikanth, N. Dey, A.N.J. Raj, A.E. Hassani, K. Santosh, N. Raja, Harmony-search and otsu based system for coronavirus disease (covid-19) detection using lung ct scan images, *arXiv preprint arXiv* (2020), 200403431.
- [24] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, et al., Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound, *IEEE Trans. Med. Imag.* (2020).
- [25] Y. Oh, S. Park, J.C. Ye, Deep learning covid-19 features on cxr using limited training data sets, *IEEE Trans. Med. Imag.* (2020).
- [26] Y.H. Wu, S.H. Gao, J. Mei, J. Xu, D.P. Fan, R.G. Zhang, et al., Jcs: an explainable covid-19 diagnosis system by joint classification and segmentation, *IEEE Trans. Image Process.* 30 (2021) 3113–3126.
- [27] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, et al., Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography, *Cell* (2020).
- [28] M. Goncharov, M. Pisov, A. Shevtsov, B. Shirokikh, A. Kurmukov, I. Blokhin, et al., Ct-based covid-19 triage: deep multitask learning improves joint identification and severity quantification, *Med. Image Anal.* 71 (2021), 102054.

- [29] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
- [30] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, et al., Attention u-net: learning where to look for the pancreas, arXiv preprint arXiv (2018), 180403999.
- [31] H. Hu, Z. Zhang, Z. Xie, S. Lin, Local relation networks for image recognition, in: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 3464–3473.
- [32] Y. Hu, Y. Chen, X. Li, J. Feng, Dynamic feature fusion for semantic edge detection, arXiv preprint arXiv (2019), 190209104.
- [33] J. Zhao, J. Liu, D. Fan, Y. Cao, J. Yang, M. Cheng, Egnet: edge guidance network for salient object detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 8779–8788.
- [34] Z. Wu, L. Su, Q.m. Huang, Stacked cross refinement network for edge-aware salient object detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 7264–7273.
- [35] Z. Zhang, H. Fu, H. Dai, J. Shen, Y. Pang, L. Shao, Et-net: A generic edge-attention guidance network for medical image segmentation, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 442–450.
- [36] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv (2015), 150203167.
- [37] Z. Fu, Y. Zheng, H. Ye, Y. Kong, J. Yang, L. He, Edge-aware deep image deblurring, arXiv preprint arXiv (2019), 190702282.
- [38] M. Long, J. Wang, G. Ding, J. Sun, P.S. Yu, Transfer feature learning with joint distribution adaptation, in: Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 2200–2207.
- [39] S.J. Pan, I.W. Tsang, J.T. Kwok, Q. Yang, Domain adaptation via transfer component analysis, IEEE Trans. Neural Network. 22 (2) (2010) 199–210.
- [40] B. Gong, Y. Shi, F. Sha, K. Grauman, Geodesic flow kernel for unsupervised domain adaptation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2012, pp. 2066–2073.
- [41] D.M. Hansell, A.A. Bankier, H. MacMahon, T.C. McLoud, N.L. Muller, J. Remy, Fleischner society: glossary of terms for thoracic imaging, Radiology 246 (3) (2008) 697–722.
- [42] S. Salehi, A. Abedi, S. Balakrishnan, A. Gholamrezaezhad, Coronavirus disease 2019 (covid-19): a systematic review of imaging findings in 919 patients, Am. J. Roentgenol. (2020) 1–7.
- [43] J. Zhao, Y. Zhang, X. He, P. Xie, Covid-ct-dataset: a ct scan dataset about covid-19, arXiv preprint arXiv (2020), 200313865.
- [44] D.P. Kingma, J. Ba, Adam, A method for stochastic optimization, arXiv preprint arXiv (2014), 14126980.
- [45] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: a deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.
- [46] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2017) 834–848.
- [47] L.C. Chen, G. Papandreou, F. Schroff, H. Adam, Rethinking atrous convolution for semantic image segmentation, arXiv preprint arXiv (2017), 170605587.
- [48] M.Z. Alom, M. Hasan, C. Yakopcic, T.M. Taha, V.K. Asari, Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation, arXiv preprint arXiv (2018), 180206955.
- [49] G. Wang, X. Liu, C. Li, Z. Xu, J. Ruan, H. Zhu, et al., A noise-robust framework for automatic segmentation of covid-19 pneumonia lesions from ct images, IEEE Trans. Med. Imag. 39 (8) (2020) 2653–2663.
- [50] M. Tan, Q.V. Le, Efficientnet: rethinking model scaling for convolutional neural networks, arXiv preprint arXiv (2019), 190511946.
- [51] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.
- [52] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1492–1500.
- [53] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, Automated detection of covid-19 cases using deep neural networks with x-ray images, Comput. Biol. Med. (2020), 103792.
- [54] C. Jin, W. Chen, Y. Cao, Z. Xu, Z. Tan, X. Zhang, et al., Development and evaluation of an artificial intelligence system for covid-19 diagnosis, Nat. Commun. 11 (1) (2020) 1–14.
- [55] H. Chen, Y. Wang, C. Xu, B. Shi, C. Xu, Q. Tian, et al., Addernet: do we really need multiplications in deep learning?, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 1468–1477.
- [56] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, C. Xu, Ghostnet: more features from cheap operations, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 1580–1589.
- [57] S. Gao, M.M. Cheng, K. Zhao, X.Y. Zhang, M.H. Yang, P.H. Torr, Res2net: a new multi-scale backbone architecture, IEEE Trans. Pattern Anal. Mach. Intell. (2019).