Multi-center sparse learning and decision fusion for automatic COVID-19 diagnosis

Zhongwei Huang [a], Haijun Lei [a], Guoliang Chen [a], Haimei Li [c], Chuandong Li [d], Wenwen Gao [d], Yue Chen [d], Yaofa Wang [e], Haibo Xu [f], Guolin Ma [d,*], Baiying Lei [b,*]

[a] Key Laboratory of Service Computing and Applications, Guangdong Province Key Laboratory of Popular High Performance Computers, Guangdong Province Engineering Center of China-made High Performance Data Computing System, Guangdong Laboratory of Artificial-Intelligence and Cyber-Economics, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China
[b] National- Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China
[c] Department of Radiology, Fu Xing Hospital, Capital Medical University, Beijing, China
[d] Department of Radiology, China-Japan Friendship Hospital, Beijing, China
[e] Minfound Medical Systems Co., Ltd., Hangzhou, China
[f] Department of Radiology, Zhongnan Hospital of Wuhan University, Wuhan, China

ARTICLE INFO

ABSTRACT

The coronavirus disease 2019 (COVID-19) pandemic caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to a sharp increase in hospitalized patients with multi-organ disease pneumonia. Early and automatic diagnosis of COVID-19 is essential to slow down the spread of this epidemic and reduce the mortality of patients infected with SARS-CoV-2. In this paper, we propose a joint multi-center sparse learning (MCSL) and decision fusion scheme exploiting chest CT images for automatic COVID-19 diagnosis. Specifically, considering the inconsistency of data in multiple centers, we first convert CT images into histogram of oriented gradient (HOG) images to reduce the structural differences between multi-center data and enhance the generalization performance. We then exploit a 3-dimensional convolutional neural network (3D-CNN) model to learn the useful information between and within 3D HOG image slices and extract multi-center features. Furthermore, we employ the proposed MCSL method that learns the intrinsic structure between multiple centers and within each center, which selects discriminative features to jointly train multi-center classifiers. Finally, we fuse these decisions made by these classifiers. Extensive experiments are performed on chest CT images from five centers to validate the effectiveness of the proposed method. The results demonstrate that the proposed method can improve COVID-19 diagnosis performance and outperform the state-of-the-art methods.

1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to a sharp increase in hospitalized patients with multi-organ disease pneumonia. In the early stages of the disease, patients with COVID-19 may be asymptomatic or appear some common symptoms, such as fever and cough [1]. With the disease worsening, these patients can appear viral sepsis, which may further lead to life-threatening organ dysfunction and even failure [2]. According to the data from the world health organization (WHO), COVID-19 has caused more than 239.4 million confirmed cases and 4.8 million deaths worldwide as of 15 October 2021 [3]. Since it spreads rapidly mainly through respiratory droplets from face-to-face contact, infection cases are increasing with a mortality rate of up to 2%, which causes the clinicians' shortage and the increase of their workloads. Many clinical measures have been used to diagnose whether the suspected cases are infected with SARS-CoV-2, such as reverse transcription-polymerase chain reaction testing (RT-PCR) [4,5], manual chest X-ray [2,6], and computed tomography (CT) [7,8] screening. However, RT-PCR easily has a low sensitivity problem [9], and the limited RT-PCR test kits cannot detect all the suspected cases. Meanwhile, manual X-ray and CT screening are time-consuming and cannot quickly and accurately diagnose COVID-19.

* Corresponding authors.
   E-mail addresses: 1950271002@email.szu.edu.cn (Z. Huang), lhj@szu.edu.cn (H. Lei), glchen@szu.edu.cn (G. Chen), 1043652709@qq.com (H. Li), lichuandong01@163.com (C. Li), 1196715172@qq.com (W. Gao), 71973292@qq.com (Y. Chen), mingfeng@zjmingfeng.com (Y. Wang), xuhaibo1120@hotmail.com (H. Xu), maguolin1007@qq.com (G. Ma), leiby@szu.edu.cn (B. Lei).

As a good plan B, artificial intelligence methods based on medical image data are playing an increasing role in automatic COVID-19 diagnosis [10–12]. In the early diagnosis of the disease, chest CT images can better reflect lung abnormalities than chest X-ray images and have captured widespread attention [2]. For example, Li et al. [10] first used RestNet50 as the backbone to extract features of each slice from CT images and then fused these features by a max-pooling operation. The final fused features were input to a fully connected layer and softmax activation function for pneumonia diagnosis. Wang et al. [13] proposed a transfer learning model based on the inception network, where the pre-trained inception network is used to extract features from CT slices, and the fully connected network is exploited for COVID-19 diagnosis. However, most of these methods treat 3D CT images as a series of slices [10,14] or manually select several pathological slices [15,16] and input them into a 2-dimensional convolutional neural network (2D-CNN) model for COVID-19 diagnosis. Therefore, they are inherently unable to exploit the context of adjacent slices to improve classification performance. Also, the manual selection of some slices consumes the doctor's time. Since the 3-dimensional CNN (3D-CNN) model can effectively and automatically learn the inter-slice context information, it has been widely used for classification and segmentation tasks of 3D data [17].

Due to the large-scale outbreak of COVID-19 around the world, the data studied in the literature are collected from different centers or hospitals, such as two centers [18], five centers [19], six centers [10,20], seven centers [21,22], and ten centers [14]. These multi-center data acquired from different devices and parameter settings are structurally inconsistent, and the lesion characteristics of some center data are not obvious. For example, in this paper, we have collected chest CT images of COVID-19 patients from Keting hospital (KT1), Wuhan shelter hospital (WH), another Keting hospital (KT2), Wuhan No. 7 Hospital (SH), and Zhongnan hospital (ZN) of Wuhan University, China, respectively. We also collect chest CT images of their corresponding normal control (NC). As shown in Fig. 1, we can see that CT images in SH and ZN centers are quite different from the other three centers, and CT images of COVID-19 patients in KT center are similar to those of the NC group, which has undoubtedly augmented the difficulties of COVID-19 diagnosis. Since the histogram of oriented gradient (HOG) representation can capture gradient structure with local shape characteristics, it is easy to control the invariance of local geometry and photometric conversion [23,24]. Inspired by the above, considering the inconsistency of data in multiple centers, we first convert 3D CT images into HOG images, which can reduce the structural differences between multi-center data. We then design a 3D-CNN model to learn the useful information between and within 3D HOG image slices and extract multi-center features, which can also save the doctor's time to screen the pathological slices.

Furthermore, most existing methods are restricted to treat multiple centers into one center [21,25], which ignores the relationship between and within multi-center data. There are some methods proposed to learn the relationship of multi-center data. For example, in [26], Wang et al. proposed a united learning scheme to facilitate the diagnosis of COVID-19 by learning heterogeneous information from different centers. In our method, considering the inconsistency of data in multiple centers, we convert 3D CT images into HOG images to reduce the structural differences between multi-center data, thereby enhancing generalization performance. These two methods solve the problem of multi-center data differences from two aspects, which form an effective complementarity. In [27], Song et al. first used a 3D-CNN to extract features from multi-center CT images and then constructed an augmented multi-center graph to consider the heterogeneity of multi-center data and disease state information.

Finally, the multi-center graph is input into a graph convolutional network (GCN) for COVID-19 diagnosis. In the process of constructing the multi-center graph, this method only constructs edges for samples with the same category at the same center, that is, it only considers the relationship within each center, while ignoring the relationship between multi-center data. Also, this method inputs the constructed graph directly into the GCN for classification and does not consider the multi-center fusion strategy to enhance the classification performance. There is redundancy in the relatively high-dimensional features extracted from CNN models, which can negatively affect the generalization ability of the model [28,29]. Besides, since information fusion can gather all kinds of useful information to enhance generalization performance, various fusion strategies (e.g., center fusion, feature fusion, and decision fusion) have received wide attention from researchers [11,30–32]. Center fusion treats multiple centers into one center [21]. Feature fusion concatenates/encodes multi-modal features into high-dimensional/another form of features [33]. Decision fusion combines the decisions made by multiple data [34]. However, center fusion can eliminate the relationship between multi-center data, and feature fusion can result in high feature dimensionality, which can reduce model generalization performance. For decision fusion, it can comprehensively consider the relationship between multi-center data and jointly analyze the diagnosis results of different classifiers to improve the diagnostic performance. Inspired by the above, we devise a multi-center sparse learning (MCSL) method to learn the intrinsic structure information between multiple centers and within each center for COVID-19 diagnosis. Specifically, we use the group sparseness regularizer to learn the intrinsic relationship between multiple centers and the global sparseness regularizer to enhance the generalization performance. Furthermore, we consider the similarity among subjects of each center to preserve the local structure information within each center. Finally, we select discriminative features from our MCSL model to jointly learn multiple classifiers. The decision is determined by fusing the results of multi-center classifiers.

We highlight the main contributions of our paper as follows.

- Considering the inconsistency of data in multiple centers, we convert 3D CT images into HOG images to reduce the structural differences between multi-center data and enhance the generalization performance.
- We propose a 3D-CNN model to learn the useful information between and within 3D HOG image slices and extract multi-center features, which saves the doctor's time to screen the pathological slices.
- We propose a MCSL method to learn the intrinsic structure information between multiple centers and within each center and select discriminative features to jointly train multi-center classifiers.
- The decision fusion is used to comprehensively consider the relationship between multi-center data and jointly analyze the diagnosis results of different classifiers to improve the diagnostic performance.

2. Related work

2.1. Feature learning

Feature learning methods aim to automatically learn deep features to improve the diagnostic performance of COVID-19. However, due to the differences between multi-center data, it is challenging to use a deep learning model to learn discriminative features that meet the accuracy requirements of multiple centers. Researchers have made some attempts by using different CNN
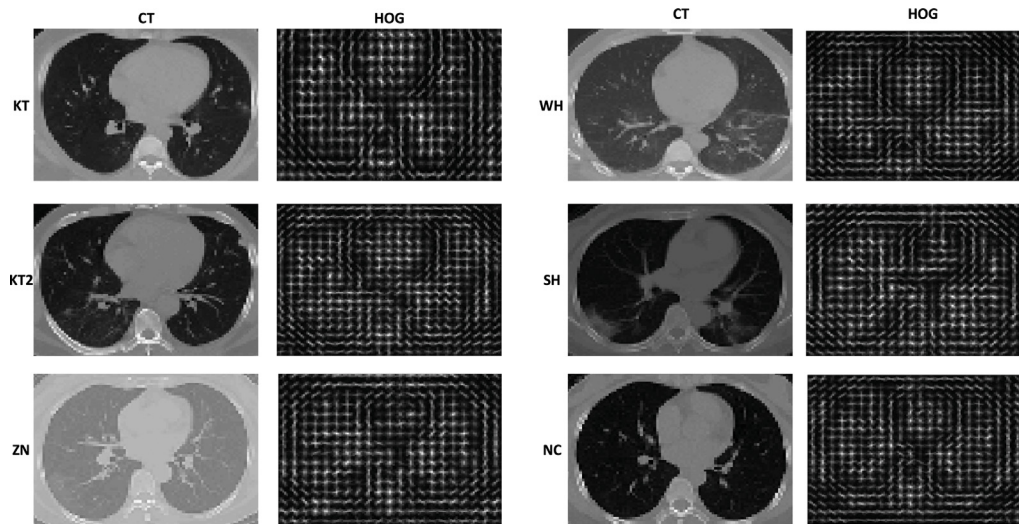
Fig. 1. Examples of CT images and corresponding HOG images in COVID-19 patients and NC.

models to handle this challenge. Ghoshal et al. [18] proposed a Bayesian convolutional neural network (BCNN) to estimate network uncertainty to improve the accuracy of pneumonia screening. Li et al. [10] first used RestNet50 as the backbone to extract features of each slice from CT images and then fused these features by a max-pooling operation. The final fused features were input to a fully connected layer and softmax activation function for pneumonia diagnosis. Gao et al. [14] used the segmented lesion features to facilitate classification at the slice level and then used a slice probability mapping method to reduce the potential influence of different imaging parameters from individual facilities for COVID-19 diagnosis. Jie et al. [22] proposed an artificial intelligence system, where CT images were divided into different cohorts to train the deep learning-based model for COVID-19 diagnosis. Wang et al. [21] developed a fully automatic deep learning system, where lung segmentation network, non-lung area suppression operation, and lung-ROI normalization were used to obtain standard lung region, and then the standardized lung-ROI was input into the COVID-19Net for COVID-19 diagnosis.

These methods show great potential by using various models to learn deep features for COVID-19 diagnosis. However, they do not fully consider processing at the data level to reduce the differences between multi-center data. Meanwhile, most of these methods treat 3D images as a series of slices or manually select several pathological slices and then input these slices into 2D-CNN models for COVID-19 diagnosis. This ignores related information between slices, and manual selection of some slices also consumes the doctor's time. In this paper, considering the inconsistency of data in multiple centers, we first convert the original CT images into HOG images, which can reduce the structural differences between multi-center data. We then propose a 3D-CNN model to automatically learn the useful information between and within 3D HOG image slices and extract multi-center features, which can also save the doctor's time to screen the pathological slices.

2.2. Sparse learning

Sparse learning methods aim to select discriminative features for various tasks such as classification, regression, and prediction. For example, in [35], the least absolute shrinkage and selection operator method (Lasso) method used linear regression and l_1 penalty to perform feature selection. In [36], the elastic net method used linear regression, l_1 norm and l_2 penalties for feature

selection. Zhang et al. [37] proposed a multi-modal multi-task method to select common relevant features for Alzheimer's disease diagnosis. Sun et al. [20] first extracted the location-specific handcrafted features from chest CT images and then used the adaptive feature selection guided deep forest (AFS-DF) model for COVID-19 classification. Shaban et al. [38] proposed a COVID-19 detection strategy based on a hybrid sparse learning and enhanced k-nearest neighbor classifier. In [39], Zhu et al. first extracted handcrafted features from CT images and then used joint logistic and linear regression to select the discriminative features for COVID-19 classification and prediction.

These methods have good performance in different tasks. However, these methods do not consider the relationship between and within multi-center or multi-modal data at the same time. In this paper, to handle this problem, we propose a MCSL method to capture the intrinsic structure between multiple centers and within each center and selects discriminative features to jointly train multi-center classifiers.

2.3. Information fusion

Information fusion aims to gather all kinds of useful information to enhance generalization performance [11,30,40,41]. There are currently three main fusion strategies, i.e., center fusion, feature fusion, and decision fusion. Center fusion treats multiple centers into one center [21]. Feature fusion concatenates/encodes multiple modal features into higher-dimensional/another form of features [33]. For example, Zhu et al. [42] first linearly fused the multi-modal features and then input the fused feature into a sparse learning model to select the discriminative features for Alzheimer's disease diagnosis. Decision fusion combines the decisions made by multiple data [34]. For example, Xie et al. [30] used decision fusion to combine texture, shape, and deep model features for the classification of lung nodules.

These fusion strategies have good performance in different scenarios. However, center fusion can eliminate the relationship between multi-center data, and feature fusion can result in high feature dimensionality, which can reduce model generalization performance. For decision fusion, it can comprehensively consider the relationship between multi-center data and jointly analyzes the diagnosis results of different classifiers to improve the diagnostic performance. Thus, in this paper, we use decision fusion to obtain the final result.
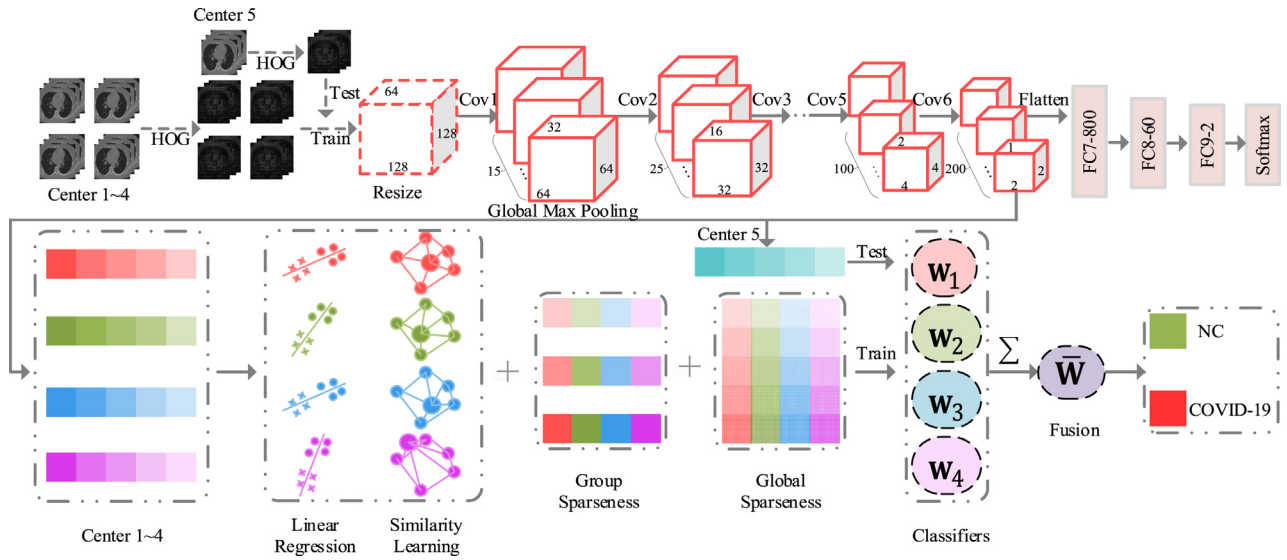
Fig. 2. Illustration of the proposed method for COVID-19 diagnosis.

3. Method

[Fig. 2](#) shows our method for automatic COVID-19 diagnosis. Specifically, considering the inconsistency of data in multiple centers, we first convert 3D CT images into HOG images, which can reduce the structural differences between multi-center data. We then use a 3D-CNN to learn deep features from each center. Finally, we employ the proposed MCSL model to learn the intrinsic structure between multiple centers and within each center and select discriminative features for COVID-19 diagnosis.

3.1. Notations

In this paper, boldface uppercase typefaces indicate matrices, boldface lowercase typefaces indicate vectors, and normal italic typefaces indicate scalars. For a matrix $\mathbf{X}$, we use $\mathbf{X}^T$, $tr(\mathbf{X})$, $\|\mathbf{X}\|_{2,1} = \sum_i \|\mathbf{x}_i\|_2$, and $\|\mathbf{X}\|_F = \sqrt{\sum_i \|\mathbf{x}_i\|_2^2}$ to indicate the transposition, trace, $\ell_{2,1}$-norm and Frobenius norm of $\mathbf{X}$, respectively.

3.2. Multi-center deep feature extraction

For each CT slice, the corresponding HOG image is generated through the following steps: (1) divide the CT slice into many small spatial cells of equal size; (2) compute gradients and accumulate histogram of gradient directions over the pixels of each cell; (3) conduct contrast normalization on local responses; (4) generate a pictorial rendition of HOG descriptors. As shown in [Fig. 2](#), we construct a 3D-CNN to automatically learn multi-center deep features from HOG images. Specifically, the input of this network is a down-sampled HOG image with a size of $128 \times 128 \times 64$. It has six convolutional layers and three fully connected (FC) layers (i.e., FC7-800, FC8-60, and FC9-2). The size of each convolutional kernel is $3 \times 3 \times 3$ and its stride, padding, and data format equal one, 'same', 'channel first', respectively. The number of convolutional kernels in each convolutional layer is 15, 25, 50, 50, 100, and 200, respectively. Each convolution layer is followed by a rectified linear unit activation function and a max-pooling operation. The size of each max-pooling operation is $2 \times 2 \times 2$ and its stride, padding, and data format equal two, 'same', and 'channel first', respectively. After building the network model, we first use four central data to train the network model to identify whether patients have COVID-19. Finally, we pop FC9-2, FC8-60, and FC7-800 and add a global max-pooling layer to

Table 1
The architecture of our 3D-CNN used in this paper.

Layer	Kernels, channels, stride, padding, data format
Conv1	$3 \times 3 \times 3$, 15, 1, same, channel first
ReLU	
Max pooling, $2 \times 2 \times 2$, 15	
Conv2	$3 \times 3 \times 3$, 25, 1, same, channel first
ReLU	
Max pooling, $2 \times 2 \times 2$, 25	
Conv3	$3 \times 3 \times 3$, 50, 1, same, channel first
ReLU	
Max pooling, $2 \times 2 \times 2$, 50	
Conv4	$3 \times 3 \times 3$, 50, 1, same, channel first
ReLU	
Max pooling, $2 \times 2 \times 2$, 50	
Conv5	$3 \times 3 \times 3$, 100, 1, same, channel first
ReLU	
Max pooling, $2 \times 2 \times 2$, 100	
Conv6	$3 \times 3 \times 3$, 200, 1, same, channel first
ReLU	
Max pooling, $2 \times 2 \times 2$, 200	
FC, $1 \times 1 \times 1$, 800	
FC, $1 \times 1 \times 1$, 60	
FC, $1 \times 1 \times 1$, 2	
Softmax	

extract multi-center deep features. The feature dimension of each center is 200. [Table 1](#) also shows the detailed information of our 3D-CNN.

3.3. Multi-center sparse learning

After learning the multi-center features, we need to use these features to build a regression model for COVID-19 diagnosis. Generally speaking, the linear multi-center regression model can be expressed as follows:

$$\min_{\mathbf{W}} \sum_{c=1}^{C} \left\| \mathbf{y}^c - \mathbf{X}^c \mathbf{w}^c \right\|_2^2, \tag{1}$$

where $\mathbf{X}^c \in \mathbb{R}^{n \times d}$ indicates the training data of n subjects and d features in the cth center, and $\mathbf{y}^c \in \{+1, -1\}$ is the corresponding true label vector (i.e., patient with COVID-19 or NC). $\mathbf{W} = \left[\mathbf{w}^1, \mathbf{w}^2, \ldots, \mathbf{w}^C\right] \in \mathbb{R}^{d \times C}$ indicates the feature weight matrix of C center data, where $\mathbf{w}^c \in \mathbb{R}^{d \times 1}$ is the feature weight vector of the cth center. The larger the value of $\mathbf{w}^c$, the more important the corresponding feature. However, there is redundancy in relatively high-dimensional features, which can lead to the over-fitting problem. To solve this problem, we add the group sparseness and global sparseness regularizers into Eq. (1), which is indicated as:

$$\min_{\mathbf{W}} \frac{1}{2} \sum_{c=1}^{C} \left\| \mathbf{y}^c - \mathbf{X}^c \mathbf{w}^c \right\|_2^2 + \lambda_1 \|\mathbf{W}\|_{2,1} + \lambda_2 \|\mathbf{W}\|_F^2, \quad (2)$$

In multi-center data, the importance of features with the same index should be similar. Thus we use the $\ell_{2,1}$-norm to conduct the group sparseness to learn the intrinsic relationship between multiple centers. The $\ell_{2,1}$-norm can make the weight matrix have many zero rows. In other words, the features corresponding to non-zero rows in $\mathbf{W}$ will be selected for COVID-19 diagnosis. Meanwhile, we use the Frobenius norm to conduct the global sparseness to reduce the overall feature weight, which can further enhance the generalization performance of the model.

Further, considering that in each center, the smaller the Euclidean distance between two subjects, the more similar their predicted target values should be. Thus, we introduce the following formula to learn this similarity relationship among subjects in each center:

$$\frac{1}{2} \sum_{c=1}^{C} \sum_{i,j}^{N_c} s_{ij}^c \left\| \mathbf{x}_i^{c^T} \mathbf{w}^c - \mathbf{x}_j^{c^T} \mathbf{w}^c \right\|_2^2 = \sum_{c=1}^{C} \left(\mathbf{X}^c \mathbf{w}^c\right)^T \mathbf{L}^c \left(\mathbf{X}^c \mathbf{w}^c\right), \quad (3)$$

where $\mathbf{x}_i^c$ and $\mathbf{x}_j^c$ indicate the ith and jth subjects in the cth center, respectively, and s_{ij}^c is their similarity value. $\mathbf{S}^c = \left[s_{ij}^c\right]$ indicates the similarity matrix of N_c subjects in the cth center. $\mathbf{L}^c = \mathbf{D}^c - \mathbf{S}^c$ indicates the Laplacian matrix of $\mathbf{S}^c$, where $\mathbf{D}^c$ is the diagonal matrix whose ith element value on the diagonal equals to $\sum_{j=1}^{N_c} \frac{s_{ij}^c + s_{ji}^c}{2}$. Eq. (3) aims to preserve the local structure information within each center. To compute the similarity matrix $\mathbf{S}^c$ in the cth center, we use a radial basis function kernel as follows:

$$s_{ij}^c = exp\left(-\frac{\left\|\mathbf{x}_i^c - \mathbf{x}_j^c\right\|_2}{4t}\right), \quad (4)$$

where t indicates a kernel width. If $\mathbf{x}_i^c$ and $\mathbf{x}_j^c$ are among the k nearest neighbors, we use Eq. (4) to compute s_{ij}^c, otherwise, we set it to zero. Finally, we add Eq. (3) into Eq. (2) and we can get our MCSL model:

$$\min_{\mathbf{W}} \frac{1}{2} \sum_{c=1}^{C} \left\| \mathbf{y}^c - \mathbf{X}^c \mathbf{w}^c \right\|_2^2 + \lambda_1 \|\mathbf{W}\|_{2,1} + \lambda_2 \|\mathbf{W}\|_F^2$$
$$+ \lambda_3 \sum_{c=1}^{C} \left(\mathbf{X}^c \mathbf{w}^c\right)^T \mathbf{L}^c \left(\mathbf{X}^c \mathbf{w}^c\right), \quad (5)$$

where λ_1, λ_2, and λ_3 are three regularization parameters. The larger their value, the stronger the binding force of the corresponding regularization terms. By our MCSL model, we can learn the intrinsic structure between multiple centers and within each center and selects discriminative features to jointly train multi-center classifiers.

3.4. Optimization of MCSL model

Since the objective function in Eq. (5) contains the non-smooth part (i.e., $\ell_{2,1}$-norm), it is difficult to solve this function directly.

Thus, as shown in Algorithm 1, we exploit the accelerated proximal gradient method (APG) [43] to optimize it. First, we divide Eq. (5) into a smooth part $h(\mathbf{W})$ and a non-smooth part $g(\mathbf{W})$:

$$h(\mathbf{W}) = \frac{1}{2} \sum_{c=1}^{C} \left\| \mathbf{y}^c - \mathbf{X}^c \mathbf{w}^c \right\|_2^2 + \lambda_2 \|\mathbf{W}\|_F^2$$
$$+ \lambda_3 \sum_{c=1}^{C} \left(\mathbf{X}^c \mathbf{w}^c\right)^T \mathbf{L}^c \left(\mathbf{X}^c \mathbf{w}^c\right), \quad (6)$$

$$g(\mathbf{W}) = \lambda_1 \|\mathbf{W}\|_{2,1}, \quad (7)$$

Then, we introduce the following function to approximate Eq. (6):

$$\beth(\mathbf{W}, \gamma_i) = h(\gamma_i) + tr\left((\mathbf{W} - \gamma_i)^T \nabla h(\gamma_i)\right) + \frac{l}{2} \|\mathbf{W} - \gamma_i\|_F^2, \quad (8)$$

where $\nabla h(\gamma_i)$ indicates the gradient of $h(\mathbf{W})$ at the search point γ_i in the ith iteration, and we can use Algorithm 2 to compute $\nabla h(\gamma_i)$. l indicates the step length. Meanwhile, we compute γ_i by the following formula:

$$\gamma_i = \mathbf{W}_i + \theta_i \left(\mathbf{W}_i - \mathbf{W}_{i-1}\right). \quad (9)$$

where $\theta_i = (\xi_{i-1} - 1)/\xi_i$ and $\xi_i = \sqrt{1 + \left(1 + 4\xi_{i-1}^2\right)}/2$ ($\xi_0 = 0$ and $\xi_1 = 1$).

Finally, according to the APG method [43], we update $\mathbf{W}_{i+1}$ as follows:

$$\mathbf{W}_{i+1} = \underset{\mathbf{W}}{\operatorname{argmin}} \|\mathbf{W} - \mathbf{P}\|_F^2 + \frac{2}{l} g(\mathbf{W})$$
$$= \sum_{j=1}^{d} \left(\left\| \mathbf{w}_j - \mathbf{p}_j \right\|_2^2 + \frac{2\lambda_1}{l} \left\| \mathbf{w}_j \right\|_2 \right), \quad (10)$$

where $\mathbf{w}_j$ and $\mathbf{p}_j$ indicate the jth row of $\mathbf{W}$ and $\mathbf{P}$, respectively. Meanwhile, $\mathbf{P}$ is obtained by the following formula:

$$\mathbf{P} = \gamma_i - \frac{1}{l} \nabla h(\gamma_i), \quad (11)$$

According to the work in [44], we can compute each $\mathbf{w}_j$ of Eq. (10) and obtain $\mathbf{W}_{i+1} = \left[\mathbf{w}_1^T; \mathbf{w}_2^T; \ldots; \mathbf{w}_d^T\right]$:

$$\mathbf{w}_j = \begin{cases} \left(1 - \dfrac{\lambda_1}{l \left\|\mathbf{p}_j\right\|_2}\right) \mathbf{p}_j, & if \ \left\|\mathbf{p}_j\right\|_2 > \dfrac{\lambda_1}{l} \\ 0, \ otherwise \end{cases} \quad (12)$$

We dynamically update $\mathbf{W}$ until convergence using Algorithm 1 and then use the optimal $\mathbf{W}$ as multi-center classifiers.

3.5. Decision fusion

Since information fusion can gather all kinds of useful information to enhance the generalization performance, various fusion strategies (e.g., center fusion, feature fusion, and decision fusion) have received wide attention from researchers [11,30,40,41]. In this paper, we select the third method, namely decision fusion. Specifically, we first exploit Algorithm 1 to optimize the feature weight $\mathbf{W}$ of Eq. (13):

$$\underset{\mathbf{W}}{\operatorname{argmin}} \frac{1}{2} \sum_{c=1}^{C} \left\| \mathbf{y}^c - \mathbf{X}^c \mathbf{w}^c \right\|_2^2, +\lambda_1 \|\mathbf{W}\|_{2,1} + \lambda_2 \|\mathbf{W}\|_F^2$$
$$+ \lambda_3 \sum_{c=1}^{C} \left(\mathbf{X}^c \mathbf{w}^c\right)^T \mathbf{L}^c \left(\mathbf{X}^c \mathbf{w}^c\right), \quad (13)$$

Algorithm 1: Solution to Eq. (5)

Input: Data matrix $\mathbf{X}^c \in \mathbb{R}^{n \times d}$ ($c = 1,2,\ldots,C$), label vector $\mathbf{y}^c \in \{+1, -1\}$, regularization parameters $\lambda_1, \lambda_2, \lambda_3$, nearest neighbor coefficient k, and kernel width t;

Output: Feature weight matrix $\mathbf{W} = [\mathbf{w}^1, \mathbf{w}^2, \ldots, \mathbf{w}^C] \in \mathbb{R}^{d \times C}$;

1 Initialize $\mathbf{w}_0^c = \mathbf{w}_1^c = \mathbf{X}^{cT}\mathbf{y}^c$, $\xi_0 = 0$, $\xi_1 = 1$, $i = 1$, $l_0 > 0$, $o > 1$, ε, and *tol*

2 **Repeat**

3    $l = l_{i-1}$;

4   Compute the coefficient $\theta_i = (\xi_{i-1} - 1)/\xi_i$ in the *i*-th iteration;

5   Compute the search point $\boldsymbol{\gamma}_i$ in the *i*-th iteration according to Eq. (9);

6   Compute the gradient $\nabla h(\boldsymbol{\gamma}_i)$ of $h(\boldsymbol{\gamma}_i)$ at the search point $\boldsymbol{\gamma}_i$ according to Algorithm 2;

7   **Repeat**

8   Compute the intermediate variable $\mathbf{P}$ according to Eq. (11);

9   Compute the feature weight matrix $\mathbf{W}_{i+1}$ in the *i+1*-th iteration according to Eq. (10);

10   Compute the value of the objective function $h(\mathbf{W}_{i+1})$ according to Eq. (6);

11   Compute the value of the objective function $\beth(\mathbf{W}_{i+1}, \boldsymbol{\gamma}_i)$ according to Eq. (8);

12   Increase the step length $l = o\, l$;

13   **Until** $h(\mathbf{W}_{i+1}) \leq \beth(\mathbf{W}_{i+1}, \boldsymbol{\gamma}_i)$ or $\|\mathbf{W}_{i+1} - \boldsymbol{\gamma}_i\|_F^2 \leq \varepsilon$;

14   Compute the function value $f(\mathbf{W}_{i+1}) = h(\mathbf{W}_{i+1}) + g(\mathbf{W}_{i+1})$;

15    $i = i + 1$

16    $\xi_i = \sqrt{1 + (1 + 4\xi_{i-1}^2)}/2$

17 **Until** $|f(\mathbf{W}_i) - f(\mathbf{W}_{i-1})| \leq$ *tol*.

18 Obtain the feature weight matrix $\mathbf{W} = \mathbf{W}_i$.

We then use the optimal multi-center weight parameter $\mathbf{W}$ as multi-center classifiers, where each $\mathbf{w}^c$ is a central classifier like a clinician. To comprehensively consider the diagnosis results of each clinician, we introduce the following formula to fuse the decision of each center to obtain the final diagnosis result for test subjects:

$$\mathbf{y}_{predict} = sign\left(\mathbf{X}_{test}\mathbf{W}\,(0.25, 0.25, 0.25, 0.25)^T\right), \qquad (14)$$

$$sign(x) = \begin{cases} +1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0, \end{cases} \qquad (15)$$

where $\mathbf{X}_{test}$ indicates the training data, and the decision weight of each center is 0.25. To further prove the advantages of decision fusion, we attempt to conduct fusion at the center level, i.e., concatenating training subjects from multiple centers, and input these subjects to the MCSL model to learn a feature weight vector $\mathbf{w}$ for COVID-19 diagnosis. Meanwhile, we also try to perform fusion at the feature level, that is, to concatenate the features of CT images and HOG images, and then input these features to the MCSL model to learn a feature weight vector $\mathbf{w}$ for COVID-19 diagnosis.

4. Experiments

4.1. Data acquisition

We have collected a total of 1034 chest CT images of COVID-19 patients from five hospitals or centers in Wuhan, China, including 178 subjects from KT1 center, 130 subjects from WH center, 417 subjects from KT2 center, 104 subjects from SH center, and 205 subjects from ZN center. We have also collected a total of 2298 chest CT images of NC. These NC subjects are allocated to each center in proportion to the number of their corresponding

---

Algorithm 2: Compute the gradient of Eq. (6)

Input: Data matrix $\mathbf{X}^c \in \mathbb{R}^{n \times d}$ $(c = 1, 2, \ldots, C)$, current feature weight matrix $\mathbf{W} = [\mathbf{w}^1, \ \mathbf{w}^2, \ldots, \mathbf{w}^C] \in \mathbb{R}^{d \times C}$, and Laplacian matrix $\mathbf{L}^c \in \mathbb{R}^{N_c \times N_c}$;

Output: Gradient matrix $\nabla h(\mathbf{W})$.

1   Initialize $c = 1$ and set $\nabla h(\mathbf{W})$ to be an empty matrix;

2   **Repeat**

3   Compute $\nabla h(\mathbf{W}) = cat\big(2, \nabla h(\mathbf{W}), \mathbf{X}^{cT}(\mathbf{X}^c \mathbf{w}^c - \mathbf{y}^c) + 2\lambda_3 \mathbf{X}^{cT} \mathbf{L}^c \mathbf{X}^c \mathbf{w}^c\big)$, where $cat(2, \mathbf{A}, \mathbf{b})$ indicates that the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ are concatenated along the column direction;

4    $c = c + 1$;

5   **Until** $c > C$;

6   Obtain the gradient matrix $\nabla h(\mathbf{W}) = \nabla h(\mathbf{W}) + 2\lambda_2 \mathbf{W}$.

Table 2
The number of subjects in five centers.

Center	COVID-19	NC
KT1	178	395
WH	130	288
KT2	417	926
SH	104	231
ZN	205	458
Total	1034	2298

patients with COVID-19, namely, 395 NC subjects in KT1 center, 288 NC subjects in WH center, 926 NC subjects in KT2 center, 231 NC subjects in SH center, and 458 NC subjects in ZN center. In Table 2, we also show the data distribution in the five centers. In this paper, all the data we used have passed the corresponding ethical approval, and the use of these data has been approved.

4.2. Experimental setting

As for data preprocessing, we first down-sample CT images to a size of $128 \times 128 \times 64$. We then exploit the *VLfeat* tool package (https://www.vlfeat.org/index.html) to generate HOG images from the down-sampled CT images. Finally, we down-sample these HOG images to a size of $128 \times 128 \times 64$ again. These preprocessing procedures can not only ensure that the generated HOG images can be normally input to the 3D-CNN network for training, but also ensure the clarity of these images.

In this paper, we collect chest CT images from five centers. We use four central data as the training set and the remaining one as the test set. In this way, we can get the classification results of each center as the test set. In the 3D-CNN model, we set the following parameters: the learning rate is 0.0001, the optimizer is Adam, the loss is categorical cross-entropy, the number of epochs is 15, and the batch size is 16. In Eq. (4), we set t equal to 1 and $k \in \{1, \ldots, 5\}$. We also tune the regularization parameters of the MCSL method as follows: $\lambda_1 \in \{10^{-2}, \ldots, 10^5\}$, λ_2 and $\lambda_3 \in \{10^{-5}, \ldots, 10^2\}$. Finally, we also show the running environment of our framework to facilitate researchers to reproduce our experiments in Table 3. It is worth noting that we use the Transplant

package of python to call the MATLAB program, and then we use MATLAB to generate HOG images and optimize the MCSL model in the paper. We compare our MCSL method with state-of-the-art multi-center sparse learning methods including: (1) the logistic regression (LogisticR) method is a multi-center learning method consisting of the logistic regression and $\ell_{2,1}$-norm in the MALSAR1.1 package [45]. (2) The LogisticR-G method is obtained by adding a global sparse norm to the LogisticR method. (3) The least-square regression (LeastR) method is a special case of our proposed MCSL method by setting $\lambda_2 = \lambda_3 = 0$. This method does not have the subject similarity norm and the global sparse norm. (4) The LeastR-G method is also a special case of our method by setting $\lambda_3 = 0$. This method does not have the subject similarity norm. (5) A library for support vector machines (LIBSVM) [46] is a widely used classifier without feature selection. (6) We directly use the proposed 3D-CNN for COVID-19 diagnosis.

We evaluate our method on the five center datasets and the experimental results are given in Table 4, Figs. 3, and 4. To estimate its diagnostic performances, we use the quantitative metrics of accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PRE), unweighted average recall (UAR), F1-score, and area under the receiver operating characteristic (ROC) curve (AUC).

4.3. Diagnostic performance for COVID-19

Table 4 shows the sensitivity results of the competing methods for different input images in different centers. Based on the above results, we can make the following findings. Compared with using CT images, we can get better and robust performance by using corresponding HOG images. For example, the diagnosis of WH and KT2 centers is relatively easy, since images of the two centers have small differences and their lesions are relatively easy to distinguish. In KT center, we can see that chest CT images are directly input to our 3D-CNN model and have low sensitivity since the lesions are not very clear. As shown in Fig. 1, the diagnosis on KT, SH, and ZN centers is difficult, since CT images with COVID-19 patients in KT center are similar to those of the NC group, and CT images in SH and ZN centers are different from the other 3 centers. However, we can easily see that HOG images are input to our 3D-CNN model and always achieve good performance. For

Table 3
The running environment of our framework.

Server	Information	Number/size/version	Remark
Operating system	Ubuntu	18.04.4	Version
Hardware	CPU	48	Number
	GPU	2/12GB/TITAN X	Number/size/version
	Memory	120 GB	Size
	Data disk	4.4T	Size
Software/package	CUDA	10.0.130	Version
	cuDNN	7.6.3	Version
	Python	3.6.10	Version
	Transplant	0.8.10	Version (Call MATLAB)
	Tensorflow-gpu	2.0.0	Version
	Keras	2.3.1	Version (3D-CNN)
	MATLAB	R2017b	Version (MCSL)
	VLfeat	0.9.21	Version

example, there are sensitivities of 86.52%, 93.08%, 99.28%, 94.23%, and 80.00% in KT, WH, KT2, SH, and ZN centers, respectively. Because there are differences between multi-center CT images, and HOG images can maintain good invariance to image geometric and optical deformations. We convert 3D CT images into HOG images, which can reduce the structural differences between multi-center data and thus improve the diagnostic performance of COVID-19.

Meanwhile, in HOG images, compared with the LogisticR, LogisticR-G, LeastR, and LeastR-G methods that do not consider the relationship between and within multi-center data at the same time, our MCSL method uses multiple regularizers to learn the relationship between and within multiple centers and thus has achieved higher performance. For example, our MCSL method has the highest sensitivity and achieves results of 94.38%, 93.08%, 99.76%, 99.04%, and 93.17% in KT, WH, KT2, SH, and ZN centers, respectively, which also has higher sensitivity than RT-PCR such as sputum (72%) and nasal swabs (63%) [47]. In addition, although there are structural differences between multi-center CT images, our MCSL method always maintains a sensitivity of up to 90% in all centers. The reason is that our method contains the sample similarity constraint, which can learn the local spatial structure of each center and capture the heterogeneous information of multiple centers. Meanwhile, it includes the group sparsity and global sparsity constraints, which can learn the common features of multiple centers and enhance the generalization ability. Therefore, our method can adapt to multi-center changes and maintain good performance. However, the LogisticR, LogisticR-G, LeastR, and LeastR-G methods cannot adapt to changes in multi-center CT images, and hence their sensitivities are low in KT, SH, and ZN centers. Finally, the ROC curves of the competing methods are compared using CT images and HOG images in Fig. 3. Fig. 4 also shows the performance of the competing methods vividly on radar charts. We can see that our method has more robust performance than other competing methods.

4.4. The influence of regularization parameters of MCSL on performance

From Table 4, we get that using HOG images has more robust performance, and our MCSL method has the best performance in HOG images. Further, to study the influence of each regularization term of our MCSL method on the classification performance, we use HOG images with better performance for ablation experiments on our MCSL method. Table 5 shows the sensitivity results of our MCSL method using HOG images under different settings of regularization parameters. From this table, we can obtain the following findings. First, when setting $\lambda_1, \lambda_2, \lambda_3 = 0$, the MCSL method only contains a multi-center least square regression term without regularization terms. We can see that its diagnostic performance is acceptable in KT2 and SH centers, but poor in KT,

Table 4
Sensitivity results of the competing methods for different input images in different centers.

Image	Method	SEN (%)				
		KT	WH	KT2	SH	ZN
CT	SVM	79.78	96.92	**96.16**	50.00	93.17
	LogisticR	41.57	92.31	70.98	19.23	15.61
	LogisticR-G	45.51	93.85	74.34	18.27	47.80
	LeastR	41.01	96.15	93.05	27.88	13.17
	LeastR-G	84.27	97.69	93.05	58.65	27.80
	3D-CNN	57.87	96.92	94.24	19.23	9.27
	MCSL	**91.57**	97.69	94.72	**93.27**	93.17
HOG	SVM	84.27	**94.62**	98.80	96.15	88.29
	LogisticR	73.03	87.69	98.80	87.50	84.39
	LogisticR-G	73.60	91.54	98.56	90.38	80.98
	LeastR	87.08	90.77	99.04	90.38	72.68
	LeastR-G	87.08	90.77	99.28	97.12	92.20
	3D-CNN	86.52	93.08	99.28	94.23	80.00
	MCSL	**94.38**	93.08	**99.76**	**99.04**	**93.17**

Table 5
Sensitivity results of our MCSL method using HOG images under different regularization terms.

Image	Hyper-parameter	SEN (%)				
		KT	WH	KT2	SH	ZN
HOG	$\lambda_1, \lambda_2, \lambda_3 = 0$	30.34	20.77	99.76	90.38	57.56
	$\lambda_1 \neq 0$	87.08	90.77	99.04	90.38	72.68
	$\lambda_2 \neq 0$	87.08	90.77	99.76	97.12	92.20
	$\lambda_3 \neq 0$	87.64	88.46	99.52	95.19	81.46
	$\lambda_1 \neq 0, \lambda_2 \neq 0$	87.08	90.77	99.28	97.12	92.20
	$\lambda_1 \neq 0, \lambda_3 \neq 0$	90.45	88.46	99.76	99.04	84.88
	$\lambda_2 \neq 0, \lambda_3 \neq 0$	89.89	**96.92**	99.28	99.04	92.20
	$\lambda_1, \lambda_2, \lambda_3 \neq 0$	**94.38**	93.08	**99.76**	**99.04**	**93.17**

WH, and ZN centers. For example, the MCSL method without regularization term obtains a sensitivity of 90.38% in SH center, but only has a sensitivity of 30.34% in KT center.

Second, when gradually adding regularization terms to our method, we can see that its performance is improved. For example, in KT center, the MCSL method including global sparseness (i.e., $\lambda_2 \neq 0$) has a sensitivity of 87.08%. The MCSL method including global sparseness and similarity constraint (i.e., $\lambda_2 \neq 0$ and $\lambda_3 \neq 0$) obtains a sensitivity of 89.89%. The MCSL method including group sparseness, global sparseness, and similarity constraint (i.e., $\lambda_1, \lambda_2, \lambda_3 \neq 0$) obtains a sensitivity of 94.38%, which shows that each regularization term of our method can affect the performance improvement.
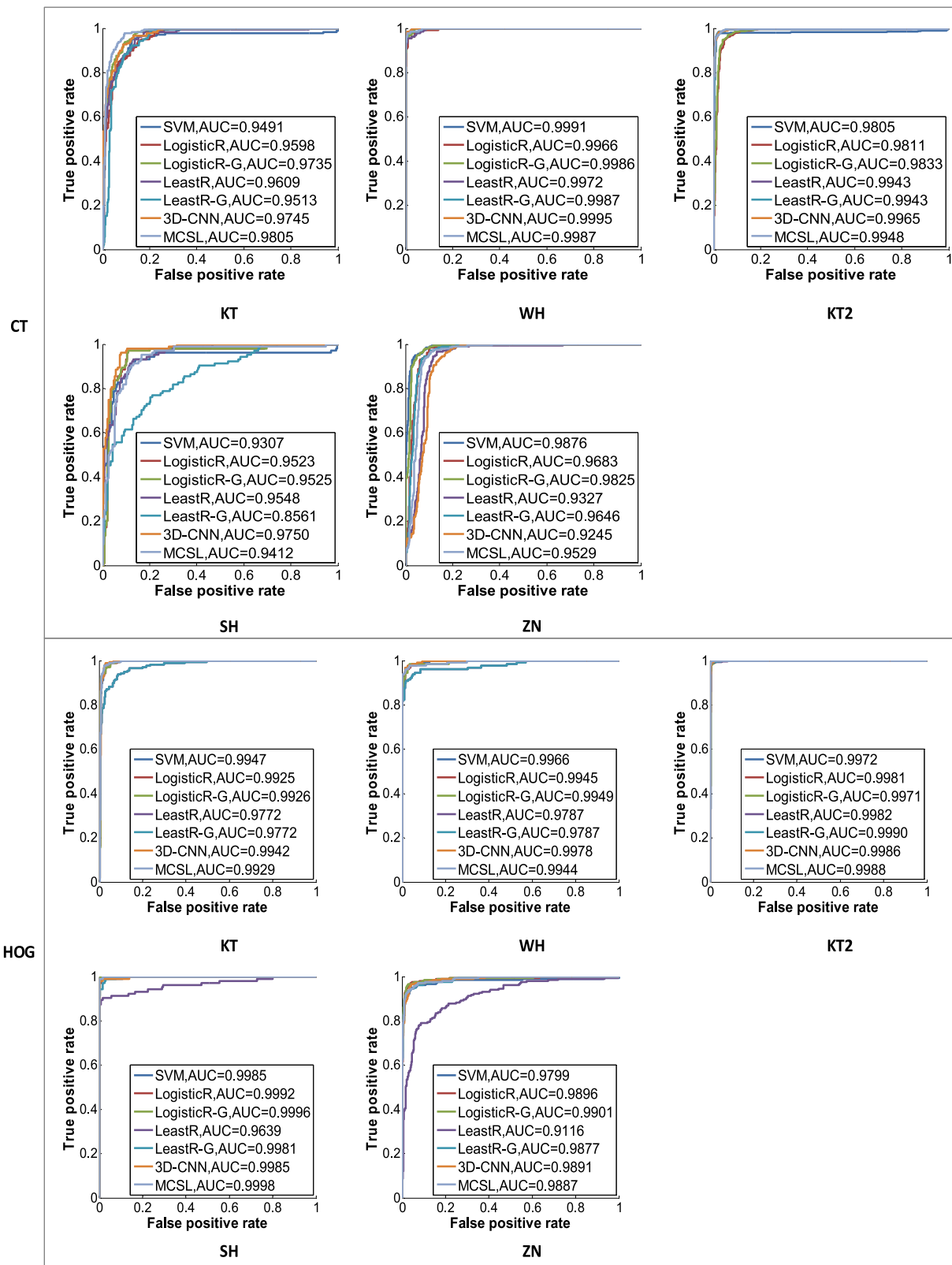
Fig. 3. ROC curves for the competing methods in CT images and HOG images.

4.5. Fusion strategies

For multi-center data, information fusion can occur at the center, feature, or decision level. In this paper, we first use 3D-CNN to extract deep features from multi-center data using HOG images. We then input these multi-center features to the MCSL model to jointly learn the feature weight matrix **W**, where each $\mathbf{w}^c$ is denoted as a clinician. Finally, to comprehensively consider the diagnosis results of each clinician, we fuse the decision of
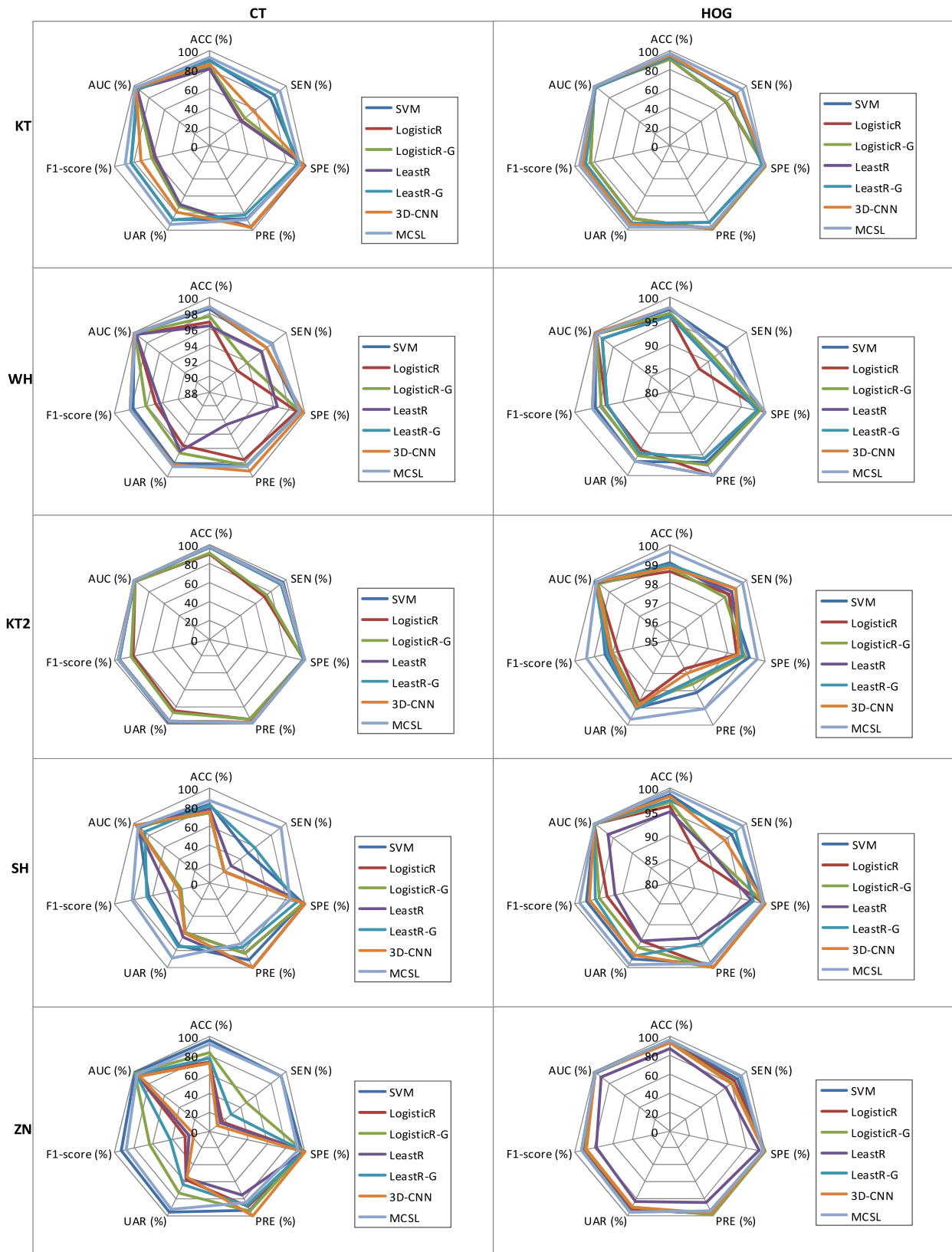
Fig. 4. Diagnostic performance of the competing methods on radar charts.

each center to obtain the final diagnosis result for test sub-jects, namely, $y_{predict} = \text{sign}\left(\mathbf{X}_{test}\mathbf{W}(0.25, 0.25, 0.25, 0.25)^T\right)$. Meanwhile, we attempt to conduct fusion at the center level, i.e., concatenating training subjects from multiple centers, and

input these subjects to the MCSL model to learn a feature weight vector **w** for COVID-19 diagnosis. We also attempt to perform fu-sion at the feature level, that is, to concatenate the features of CT images and HOG images, and then input these linearly connected

Table 6
Diagnosis performance of center fusion, feature fusion, and decision fusion.

Center	Method	ACC (%)	SEN (%)	SPE (%)	PRE (%)	UAR (%)	F1-score (%)	AUC (%)
KT	Center fusion	96.51	**94.94**	97.22	93.89	96.08	94.41	98.79
	Feature fusion	96.51	91.01	**98.99**	97.59	95.00	94.19	99.24
	Decision fusion	**97.21**	94.38	98.48	**96.55**	**96.43**	**95.45**	**99.29**
WH	Center fusion	96.41	93.08	97.92	95.28	95.50	94.16	98.75
	Feature fusion	94.74	**98.46**	93.06	86.49	95.76	92.09	99.38
	Decision fusion	**97.85**	93.08	**100.00**	**100.00**	**96.54**	**96.41**	**99.44**
KT2	Center fusion	99.18	99.28	99.14	98.10	99.21	98.69	**99.94**
	Feature fusion	98.88	99.04	98.81	97.41	98.93	98.22	99.57
	Decision fusion	**99.63**	**99.76**	**99.57**	**99.05**	**99.66**	**99.40**	99.88
SH	Center fusion	92.24	87.50	94.37	87.50	90.94	87.50	96.29
	Feature fusion	95.52	94.23	96.10	91.59	95.17	92.89	97.78
	Decision fusion	**99.40**	**99.04**	**99.57**	**99.04**	**99.30**	**99.04**	**99.98**
ZN	Center fusion	91.86	81.95	96.29	90.81	89.12	86.15	96.77
	Feature fusion	93.36	92.68	93.67	86.76	93.18	89.62	97.39
	Decision fusion	**96.08**	**93.17**	**97.38**	**94.09**	**95.28**	**93.63**	**98.87**

features to the MCSL model to learn a feature weight vector **w** for COVID-19 diagnosis. In Table 6, we compare the diagnosis performance of the three fusion methods, and the results show that decision fusion has better performance. For example, decision fusion, center fusion, and feature fusion methods have accuracies of 97.21%, 96.51%, and 96.51% in KT center, accuracies of 97.85%, 96.41%, and 94.74% in WH center, accuracies of 99.63%, 99.18%, and 98.88% in KT2 center, accuracies of 99.40%, 92.24%, and 95.52% in SH center, and accuracies of 96.08%, 91.86%, and 93.36% in ZN center. Decision fusion can comprehensively consider the relationship between multi-center data and jointly analyze the diagnosis results of different classifiers to obtain the best performance, which is very similar to the situation in real life where multiple doctors jointly diagnose suspected cases to improve the diagnosis accuracy. The disadvantage of center fusion may be that multiple centers are merged into one center, and the relationship between multi-center data is ignored, thereby eliminating the advantages of multi-center decision-making. The disadvantage of feature fusion may be that the fused features have higher dimensionality, and different features may also interfere with each other, which reduces the diagnostic performance. Meanwhile, in feature fusion, high-dimensional features take more time to train, and the time consumption is about three times that of decision fusion and central fusion, which is not conducive to fast training of the model to suit different scenarios.

5. Discussions

5.1. Comparison with the related methods for COVID-19 diagnosis

Based on HOG images corresponding to chest CT, we evaluate our MCSL methods against eleven state-of-the-art COVID-19 diagnosis methods. Table 7 shows the performance of these competing methods. We can see that although the method from Toraman et al. [48] has high sensitivity, they use limited X-ray images for the diagnosis, and hence their robustness needs to be verified. Meanwhile, most of these methods are based on CT images instead of X-ray images for imaging screening of COVID-19 since chest CT can better reflect lung abnormalities in the early diagnosis of the disease [2]. In addition, compared with CNN methods, we also can see that our 3D-CNN has the overall best performance, such as mean accuracy of 96.64%, mean sensitivity of 90.62%, and mean specificity of 99.35% in five centers. After using our 3D-CNN model to extract multi-center features, we use the MCSL method for diagnosis. We can see that diagnostic performance has been further improved. The reason is that these competing methods are restricted to a single-center or treat multiple centers into one center and thus ignore the

relationship between and within multi-center data. However, our MCSL method can select the most discriminative features by learning the relationship between and within multiple centers for COVID-19 diagnosis and thus has achieved higher performance. For example, our MCSL method has the best performance overall and obtains mean accuracy of 98.03%, mean sensitivity of 95.89%, and mean specificity of 99.00% in five centers.

5.2. Visualization analysis

We use the t-distributed stochastic neighbor embedding (t-SNE) method of MATLAB software to visualize high-level features. We pop FC9-2, FC8-60, and FC7-800 layers and add a global max-pooling layer to extract these features from testing images (e.g., CT images and HOG images). Fig. 5 shows the t-SNE visualization results of the features extracted from 3D-CNN using CT images and HOG images, respectively, where red dots indicate patients with COVID-19 and blue dots indicate NCs. From the t-SNE feature distribution maps, we can get the following findings. The features extracted from CT images have large intra-class differences and small inter-class differences. For example, in KT center, blue dots of the same category are divided into two pieces, and blue and red dots of different categories are close to each other. After we convert CT images to HOG images, we can see that the intra-class differences of the corresponding feature maps have become smaller, while the inter-class differences have become larger, which indicates that HOG images can improve the diagnostic performance.

5.3. Limitations and future direction

Our approach has achieved appealing performance, but there are still limitations, which we highlight here. First, when using 3D-CNN to extract deep features, we do not consider the relationship between multi-center data. We can add the relationship among multi-center data to this network to further learn the more discriminative features. Second, in this paper, we only use a single modality (i.e., chest CT image) for COVID-19 diagnosis. We can use multi-modal data (e.g., chest CT, chest radiograph, and clinical information) to improve diagnosis accuracy [30]. Third, our MCSL model is supervised, and we can further explore a semi-supervised method to adapt to scenarios with a large number of suspected cases [50]. Finally, we do not conduct the lung segmentation and lesion area positioning tasks in the paper [51]. Since SARS-CoV-2 mainly infects human lungs, we can perform the lung segmentation task to further improve diagnostic performance for COVID-19 in the future, and conduct the lesion area positioning tasks to assist doctors in diagnosing COVID-19 [22,52].
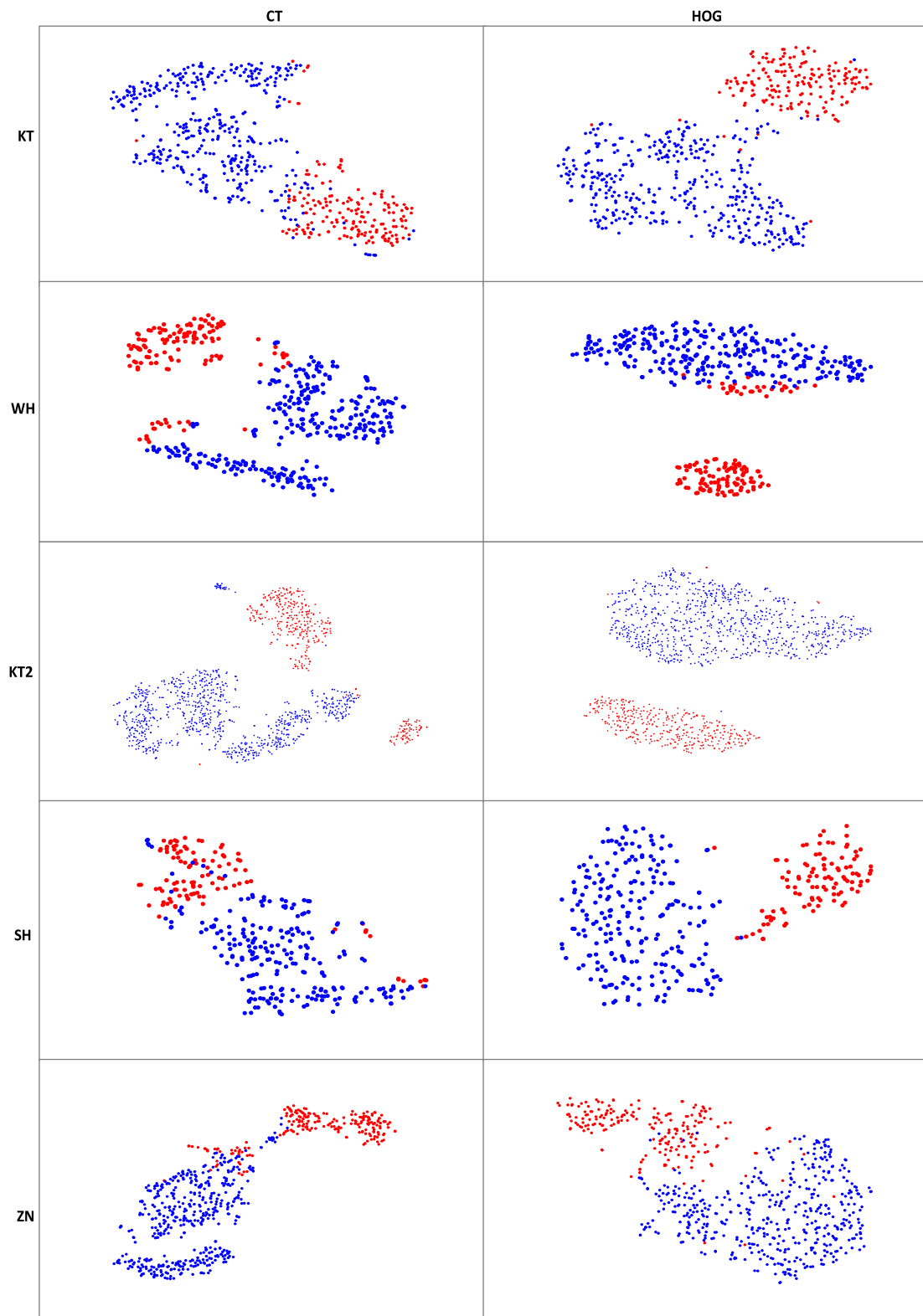
Fig. 5. The t-SNE visualization results illustrate features extracted from 3D-CNN using CT images and HOG images in five centers. (Red dots indicate patients with COVID-19 and blue dots indicate NCs).

6. Conclusions

In this paper, we propose a MCSL and decision fusion scheme exploiting chest CT images for automatic COVID-19 diagnosis. Specifically, considering the inconsistency of data in multiple centers, we first convert 3D CT images into HOG images to reduce the structural differences between multi-center data and enhance the generalization performance. We then employ a 3D-CNN model to learn the useful information between and within 3D HOG image slices and extract multi-center features, which saves the doctor's time to screen the pathological slices. Furthermore, we

Table 7
Performance comparison of some related methods for COVID-19 diagnosis (Mean).

Image	Method	Ref.	Subject	ACC	SEN	SPE
X-ray	BCNN	[18]	Two centers: 68 COVID-19, 2786 bacterial pneumonia, 1504 non-COVID-19 viral pneumonia, 1583NC	89.82	/	/
X-ray	CapsNet	[48]	One center: 231 COVID-19, 500 NC	91.24	96.00	80.95
CT	ResNet	[10]	Six centers: 468 COVID-19, 1551 CAP, 1303 NC	/	90.33	94.67
CT	U-net, DCN, FCN	[14]	Ten centers: 704 COVID-19, 498 NC	94.81	95.39	94.46
CT	DenseNet	[21]	Seven centers: 924 COVID-19, 342 other pneumonia	81.15	79.70	76.40
CT	U-net, DeCoVNet	[49]	One center: 313 COVID-19, 229 NC	90.10	84.00	98.20
CT	U-Net, ResNet	[22]	Seven centers: 3084 COVID-19, 5941 others	/	87.03	96.60
CT	U-Net++, ResNet	[19]	Five centers: 723 COVID-19, 413 NC	/	**97.40**	92.20
CT	Feature selection, KNN	[38]	Unknown: 216 COVID-19, unknown	96.00	74.00	/
CT	AFS-DF	[20]	Six centers: 1495 COVID-19, 1027 CAP	91.79	93.05	89.95
CT-H	3D-CNN	Ours	Five centers: 1034 COVID-19, 2298 NC	96.64	90.62	**99.35**
CT-H	3D-CNN, MCSL	Ours	Five centers: 1034 COVID-19, 2298 NC	**98.03**	95.89	99.00

Note: CT-H denotes the HOG images corresponding to chest CT images; boldface denotes the best performance; community-acquired pneumonia (CAP).

employ the proposed MCSL model that learns the intrinsic structure between multiple centers and within each center and selects discriminative features to jointly train multi-center classifiers. Finally, the decision fusion is used to comprehensively consider the relationship between multi-center data and jointly analyze the diagnosis results of different classifiers to improve the diagnostic performance. The extensive experiments are performed on the chest CT images from the five centers to validate the effectiveness of the proposed method, where four central data are used as the training set and the remaining one as the testing set. The results demonstrate that the proposed method can improve COVID-19 diagnosis performance and precede state-of-the-art methods.

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.

Acknowledgments

References

[1] Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, M. Wang, Presumed asymptomatic carrier transmission of COVID-19, JAMA 323 (2020) 1406–1407.

[2] W.J. Wiersinga, A. Rhodes, A.C. Cheng, S.J. Peacock, H.C. Prescott, Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (covid-19): a review, JAMA 324 (2020) 782–793.

[3] WHO, Who coronavirus disease (COVID-19) dashboard, 2021, https://covid19.who.int/ (accessed 15 2021).

[4] V.M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D.K. Chu, T. Bleicker, S. Brünink, J. Schneider, M.L. Schmidt, D.G. Mulders, B.L. Haagmans, B. van der Veer, S. van den Brink, L. Wijsman, G. Goderski, J.-L. Romette, J. Ellis, M. Zambon, M. Peiris, H. Goossens, C. Reusken, M.P. Koopmans, C. Drosten, Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR, Euro Surveill. 25 (2020) (2019) 2000045.

[5] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases, Radiology 296 (2020) E32–E40.

[6] T.B. Chandra, K. Verma, B.K. Singh, D. Jain, S.S. Netam, Coronavirus disease (COVID-19) detection in chest X-ray images using majority voting based classifier ensemble, Expert Syst. Appl. 165 (2021) 113909.

[7] J.P. Kanne, B.P. Little, J.H. Chung, B.M. Elicker, L.H. Ketai, Essentials for radiologists on COVID-19: An update-radiology scientific expert panel, Radiology 296 (2020) E113–E114.

[8] C. Long, H. Xu, Q. Shen, X. Zhang, B. Fan, C. Wang, B. Zeng, Z. Li, X. Li, H. Li, Diagnosis of the coronavirus disease (COVID-19): rRT-PCR or ct? Eur. J. Radiol. 126 (2020) 108961.

[9] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, W. Ji, Sensitivity of chest CT for COVID-19: Comparison to RT-PCR, Radiology 296 (2020) E115–E117.

[10] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, J. Xia, Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy, Radiology 296 (2020) E65–E71.

[11] S.-H. Wang, V.V. Govindaraj, J.M. Górriz, X. Zhang, Y.-D. Zhang, Covid-19 classification by fgcnet with deep feature fusion from graph convolutional network and convolutional neural network, Inf. Fusion 67 (2021) 208–229.

[12] M. Nour, Z. Cömert, K. Polat, A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization, Appl. Soft Comput. 97 (2020) 106580.

[13] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, B. Xu, A deep learning algorithm using CT images to screen for corona virus disease (COVID-19), Eur. Radiol. 31 (2021) 6096–6104.

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

[15] A. Castiglione, P. Vijayakumar, M. Nappi, S. Sadiq, M. Umer, Covid-19: Automatic detection of the novel coronavirus disease from CT images using an optimized convolutional neural network, IEEE Trans. Ind. Inf. 17 (2021) 6480–6488.

[16] S.-H. Wang, D.R. Nayak, D.S. Guttery, X. Zhang, Y.-D. Zhang, Covid-19 classification by ccshnet with deep fusion using transfer learning and discriminant correlation analysis, Inf. Fusion 68 (2021) 131–148.

[17] S. Ji, W. Xu, M. Yang, K. Yu, 3D convolutional neural networks for human action recognition, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 221–231.

[18] B. Ghoshal, A. Tucker, Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection, 2020, arXiv preprint arXiv: 2003.10769.

[19] B. Wang, S. Jin, Q. Yan, H. Xu, C. Luo, L. Wei, W. Zhao, X. Hou, W. Ma, Z. Xu, Z. Zheng, W. Sun, L. Lan, W. Zhang, X. Mu, C. Shi, Z. Wang, J. Lee, Z. Jin, M. Lin, H. Jin, L. Zhang, J. Guo, B. Zhao, Z. Ren, S. Wang, W. Xu, X. Wang, J. Wang, Z. You, J. Dong, Ai-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system, Appl. Soft Comput. 98 (2021) 106897.

[20] L. Sun, Z. Mo, F. Yan, L. Xia, F. Shan, Z. Ding, B. Song, W. Gao, W. Shao, F. Shi, H. Yuan, H. Jiang, D. Wu, Y. Wei, Y. Gao, H. Sui, D. Zhang, D. Shen, Adaptive feature selection guided deep forest for COVID-19 classification with chest CT, IEEE J. Biomed. Health Inf. 24 (2020) 2798–2805.

[21] S. Wang, Y. Zha, W. Li, Q. Wu, X. Li, M. Niu, M. Wang, X. Qiu, H. Li, H. Yu, W. Gong, Y. Bai, L. Li, Y. Zhu, L. Wang, J. Tian, A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis, Eur. Respir. J. 56 (2020) 2000775.

[22] C. Jin, W. Chen, Y. Cao, Z. Xu, Z. Tan, X. Zhang, L. Deng, C. Zheng, J. Zhou, H. Shi, J. Feng, Development and evaluation of an artificial intelligence system for COVID-19 diagnosis, Nature Commun. 11 (2020) 5088.

[23] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, pp. 886–893.

[24] S. Tian, U. Bhattacharya, S. Lu, B. Su, Q. Wang, X. Wei, Y. Lu, C.L. Tan, Multilingual scene character recognition with co-occurrence of histogram of oriented gradients, Pattern Recognit. 51 (2016) 125–134.

[25] T. Zhou, H. Lu, Z. Yang, S. Qiu, B. Huo, Y. Dong, The ensemble deep learning model for novel COVID-19 on CT images, Appl. Soft Comput. 98 (2021) 106885.

[26] Z. Wang, Q. Liu, Q. Dou, Contrastive cross-site learning with redesigned net for COVID-19 CT classification, IEEE J. Biomed. Health Inf. 24 (2020) 2806–2813.

[27] X. Song, H. Li, W. Gao, Y. Chen, T. Wang, G. Ma, B. Lei, Augmented multicenter graph convolutional network for COVID-19 diagnosis, IEEE Trans. Ind. Inf. 17 (2021) 6499–6509.

[28] H. Lei, Z. Huang, J. Zhang, Z. Yang, E.-L. Tan, F. Zhou, B. Lei, Joint detection and clinical score prediction in parkinson's disease via multi-modal sparse learning, Expert Syst. Appl. 80 (2017) 284–296.

[29] B. Lei, P. Yang, T. Wang, S. Chen, D. Ni, Relational-regularized discriminative sparse learning for alzheimer's disease diagnosis, IEEE Trans. Cybern. 47 (2017) 1102–1113.

[30] Y. Xie, J. Zhang, Y. Xia, M. Fulham, Y. Zhang, Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT, Inf. Fusion 42 (2018) 102–110.

[31] Y. Liu, C. Jiang, H. Zhao, Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums, Decis. Support Syst. 105 (2018) 1–12.

[32] Y. Wang, W. Xu, Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud, Decis. Support Syst. 105 (2018) 87–95.

[33] L.A. Alexandre, Gender recognition: A multiscale decision fusion approach, Pattern Recognit. Lett. 31 (2010) 1422–1427.

[34] H. Lei, Z. Huang, F. Zhou, A. Elazab, E. Tan, H. Li, J. Qin, B. Lei, Parkinson's disease diagnosis via joint learning from multiple modalities and relations, IEEE J. Biomed. Health Inf. 23 (2019) 1437–1449.

[35] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Statist. Soc. Ser. B-Methodol. 58 (1996) 267–288.

[36] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. R. Stat. Soc. Ser. B Stat. Methodol. 67 (2005) 301–320.

[37] D. Zhang, D. Shen, Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in alzheimer's disease, NeuroImage 59 (2012) 895–907.

[38] W.M. Shaban, A.H. Rabie, A.I. Saleh, M.A. Abo-Elsoud, A new COVID-19 patients detection strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier, Knowl.-Based Syst. 205 (2020) 106270.

[39] X. Zhu, B. Song, F. Shi, Y. Chen, R. Hu, J. Gan, W. Zhang, M. Li, L. Wang, Y. Gao, F. Shan, D. Shen, Joint prediction and time estimation of COVID-19 developing severe symptoms using chest CT scan, Med. Image Anal. 67 (2021) 101824.

[40] M. Mokhtari, H. Rabbani, A. Mehri-Dehnavi, R. Kafieh, M.-R. Akhlaghi, M. Pourazizi, L. Fang, Local comparison of cup to disc ratio in right and left eyes based on fusion of color fundus images and OCT B-scans, Inf. Fusion 51 (2019) 30–41.

[41] R. Dian, S. Li, L. Fang, Q. Wei, Multispectral and hyperspectral image fusion with spatial–spectral sparse representation, Inf. Fusion 49 (2019) 262–270.

[42] X. Zhu, H.I. Suk, D. Shen, A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis, NeuroImage 100 (2014) 91–105.

[43] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (2009) 183–202.

[44] X. Chen, W. Pan, J.T. Kwok, J.G. Carbonell, Accelerated gradient method for multi-task sparse learning problem, in: 2009 Ninth IEEE International Conference on Data Mining, 2009, pp. 746–751.

[45] J. Zhou, J. Chen, J. Ye, Malsar: Multi-Task Learning Via Structural Regularization, Vol. 21, Arizona State University, 2011, pp. 1–50.

[46] C.-C. Chang, C.-J. Lin, Libsvm: A library for support vector machines, ACM Trans. Intell. Syst. Technol. (TIST) 2 (2011) 1–27.

[47] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, W. Tan, Detection of SARS-CoV-2 in different types of clinical specimens, JAMA 323 (2020) 1843–1844.

[48] S. Toraman, T.B. Alakus, I. Turkoglu, Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks, Chaos Solitons Fractals 140 (2020) 110122.

[49] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, C. Zheng, A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, IEEE Trans. Med. Imaging 39 (2020) 2615–2625.

[50] E. Adeli, K. Thung, L. An, G. Wu, F. Shi, T. Wang, D. Shen, Semi-supervised discriminative classification robust to sample-outliers and feature-noises, IEEE Trans. Pattern Anal. Mach. Intell. 41 (2019) 515–522.

[51] Y.H. Wu, S.H. Gao, J. Mei, J. Xu, D.P. Fan, R.G. Zhang, M.M. Cheng, Jcs: An explainable COVID-19 diagnosis system by joint classification and segmentation, IEEE Trans. Image Process. 30 (2021) 3113–3126.

[52] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, Y. Shi, Abnormal lung quantification in chest CT images of COVID-19 patients with deep learning and its application to severity prediction, Med. Phys. 48 (2021) 1633–1645.