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.

Applied Soft Computing 114 (2022) 108041

Contents lists available at ScienceDirect

Applied Soft Computing

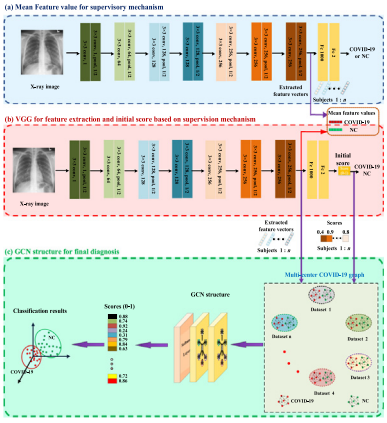
journal homepage: www.elsevier.com/locate/asoc

Novel multi-site graph convolutional network with supervision mechanism for COVID-19 diagnosis from X-ray radiographs

Ahmed Elazab [a,b], Mohamed Abd Elfattah [b], Yuexin Zhang [c,*]

[a] School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, 518060, China
[b] Computer Science Department, Misr Higher Institute for Commerce and Computers, Mansoura, Egypt
[c] School of Automation, Guangdong Polytechnic Normal University, Guangzhou, 510665, China

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 28 May 2021
Received in revised form 26 October 2021
Accepted 3 November 2021
Available online 16 November 2021

Keywords:
COVID-19
Chest X-ray radiographs
Supervision mechanism
VGG-16 network
Multi-site graph convolutional network

ABSTRACT

The novel Coronavirus disease 2019 (COVID-2019) has become a global pandemic and affected almost all aspects of our daily life. The total number of positive COVID-2019 cases has exponentially increased in the last few months due to the easy transmissibility of the virus. It can be detected using the nucleic acid test or the antibodies blood test which are not always available and take several hours to get the results. Therefore, researchers proposed computer-aided diagnosis systems using the state-of-the-art artificial intelligence techniques to learn imaging biomarkers from chest computed tomography and X-ray radiographs to effectively diagnose COVID-19. However, previous methods either adopted transfer learning from a pre-trained model on natural images or were trained on limited datasets. Either cases may lead to accuracy deficiency or overfitting. In addition, feature space suffers from noise and outliers when collecting X-ray images from multiple datasets. In this paper, we overcome the previous limitations by firstly collecting a large-scale X-ray dataset from multiple resources. Our dataset includes 11,312 images collected from 10 different data repositories. To alleviate the effect of the noise, we suppress it in the feature space of our new dataset. Secondly, we introduce a supervision mechanism and combine it with the VGG-16 network to consider the differences between the COVID-19 and healthy cases in the feature space. Thirdly, we propose a multi-site (center) COVID-19 graph convolutional network (GCN) that exploits dataset information, the status of training samples, and initial scores to effectively classify the disease status. Extensive experiments using different convolutional neural network-based methods with and without the supervision mechanism and

* Corresponding author.
E-mail address: zhangyuexin@gpnu.edu.cn (Y. Zhang).

https://doi.org/10.1016/j.asoc.2021.108041

different classifiers are performed. Results demonstrate the effectiveness of the proposed supervision mechanism in all models and superior performance with the proposed GCN.

1. Introduction

The outbreak of the ongoing coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-COV-2), urged the World Health Organization (WHO) to declare the statement of Public Health Emergency of International Concern (PHEIC) on 30th January, 2020 [1]. Later on 11th March 2020, the WHO declared COVID-19 as a global pandemic. As of 11th October, 2021, WHO reported 236,599,025 confirmed cases of COVID-19 in more than 227 countries and territories, including 4,831,486 deaths (death rate is 2.04%) [2]. Yet, the confirmed cases can be exponentially increased due to the easy transmissibility of the virus and its new strains [3], particularly, when precautions are not applied. COVID-19 patients commonly suffer from mild to severe cough, fever, and breathing problems as well as pneumonia symptoms [4]. However, asymptomatic COVID-19 cases impose a great burden on the spread of the virus as they cannot be easily detected [5]. Although research institutions and pharmaceutical entities have already developed some vaccines that could help combating COVID-19, the effectiveness of such vaccines is still arguable. Therefore, early detection and diagnosis of COVID-19 remains a paramount task in disease prevention and control.

The COVID-19 consists of single-stranded ribonucleic acid associated with 4 main structural proteins (Spike, Envelope, Membrane, and Nucleocapsid) which can be detected using either the nucleic acid test or the antibodies present in subject's blood [6]. Therefore, it was recommended by Chinese government to confirm the diagnosis of COVID-19 using the reverse transcription polymerase chain reaction (RT-PCR) [7]. However, the RT-PCR test kits availability is challenging in many regions in low- and middle-income countries that leads to difficulties in samples collection and transportation. In addition, lab tests also suffer from high false-negative cases that likely happen during samples' preparation and quality control [8]. Furthermore, the RT-PCR process takes relatively long time (4–6 h) to get results and may not be always efficient [9]. Hence, many researchers tried to introduce fast, accurate, and low-cost methods for COVID-19 diagnosis.

To this end, many researchers utilized the findings from medical imaging techniques (X-ray and computed tomography (CT)) to early diagnose the suspected COVID-19 cases and other lung diseases as well [10]. In fact, these techniques, particularly CT, have demonstrated to be an effective and fast alternative to the RT-PCR with high detection accuracy [7,11]. In addition, imaging data can be stored for disease follow-up measurements. Usually, chest X-ray or CT images are examined by 2 or more expert radiologists to determine the presence of the infection. However, due to the shortage of radiologists and the long examination time of the scans, developing computer-aided diagnosis systems becomes increasingly required for COVID-19. Despite CT scan provides richer information about lung lesions than the X-ray counterpart, it has higher cost, not always available, and takes longer acquisition time. Therefore, X-ray remains the gold standard imaging modality for diagnosis chest diseases, especially in low-income and developing countries.

Recently, many methods were proposed in the literature using the emerging technologies of artificial intelligence to, screen, detect, classify, and assess the diagnosis of COVID-19 from chest X-ray images, see the following reviews [12–18]. The breakthroughs of computer vision and deep learning methods have been effectively used in many studies to extract high level features. Majority of the methods in the literature adopted different variants of deep convolutional neural networks (CNN) [19], ResNet [20], and/or VGGNet [21] to classify the patient's image to normal, bacterial pneumonia, viral pneumonia, or COVID-19. However, many of the current methods employ transfer learning techniques to alleviate the issue of data availability for training, e.g. [22–26]. Furthermore, previous studies used insufficient number of samples to train the models and likely suffer from overfitting. Contrarily, combining samples from different hospitals/centers can improve the robustness of diagnosis system, therefore it attracts much attentions. Despite some studies combined different datasets to increase the number of training samples, the inter-dataset variations produce an undesirable noise in the combined dataset which consequently affects model's performance, if not suppressed. In addition, different hospitals all over the world usually use different imaging equipment and acquisition standards which result in an inter-variance among the data samples. Hence, alleviating this noise is highly desirable. Moreover, currents studies only consider the imaging biomarkers and ignore other useful information, that could improve diagnosis performance. For example, the statistical information of the extracted features, the interactions between samples in the classifiers, and relationship between datasets from different hospitals.

In this paper, we propose a novel method based on graph convolutional network (CGN) to incorporate other information with extracted features. The main contributions of this paper are as follow. First, we use a VGG-16 (hereinafter, we simply use VGG) network with a novel supervision mechanism to extract features for every sample in our multi-center/site datasets. To alleviate the effect of the noise in the extracted features from the different sites/centers, we propose a supervision mechanism in the training process. Specifically, after pre-training the VGG network, we get the statistical information (mean values) of the extracted features in those training samples, and then we use the statistical information as a supervision to train VGG network again. Second, we propose a multi-site COVID-19 GCN as a classifier to accomplish the diagnosis task based on the previously extracted features. In this part, we further consider the difference between datasets and design a multi-site COVID-19 graph in GCN to establish the interactions between the samples. The proposed COVID-19 graph exploits dataset information, status of training samples, and initial scores. Third, we collect a large-scale dataset from multi-centers to boost the training process and avoid the potential occurrence of overfitting. Our dataset contains 11,312 X-ray radiographs from 10 different centers/sites.

The rest of this paper is organized as follows. Section 2 discusses the related studies on COVID-19 diagnosis. In Section 3, we presents our proposed GCN method. Section 4 is dedicated for the experimental results and comparisons with state-of-the-art methods. Discussion is presented in Section 5. Finally, conclusions and perspectives are given in Section 6.

2. Related work

In this Section we revisit the major works on COVID-19 diagnosis based on the major backend architectures. Specifically, we focus on the methods that used variants of CNN-based architectures: ResNet, VGG, U-Net, and sequence-based ones.

Deep CNN architectures have been extensively used in many detection, recognition, and classification tasks. Basically, CNN extracts high level features through multiple stages and hence can effectively learn the inherent representation of the underlying data. Likewise, many studies adopted CNN-based models for COVID-19 diagnosis from both CT and X-ray images [15,22–53]. For instance, Ozturk et al. [40] trained a typical end-to-end CNN model (DarkCovidNet) that consists of 19 convolutional layers to perform binary and multiclass classification of X-ray images. This model achieved good accuracy for binary and multiclass classifications and it was evaluated by expert radiologists. In another work, Islam et al. [35] exploited CNN to extract high level feature from X-ray images and adopted long short-term memory (LSTM) network for classification. Wang and Wong [44] proposed a deep model (COVID-Net) to detect COVID-19 cases from relatively big dataset of X-ray radiographs with different disease classes. Zhang et al. [48] proposed a confidence-aware anomaly detection (CAAD) model for viral pneumonia screening to differentiate the viral pneumonia from non-viral pneumonia and healthy controls as a one-class classification-based model. The CAAD model contained an anomaly detection network and a confidence prediction network based on CNN and it was trained on large X-ray image dataset. Chowdhury et al. [30] proposed a parallel-dilated CNN-based (PDCOVIDNet) model that used a dilated convolution in the parallel stack of convolutional blocks to capture and propagate discriminative features to detect COVID-19 from X-ray radiographs.

The VGGNet architecture has been also widely adopted in many studies for COVID-19 diagnosis. For example, Panwar et al. [41] proposed an nCOVnet model that included a 24 layer CNN layers. A pre-trained VGG16 network on the ImageNet dataset was used for feature extraction in the top layers of their model and another 5 customized layers. Brunese et al. [28] proposed a three phases approach to firstly detect the presence of a pneumonia in chest radiographs then differentiate between the COVID-19 and pneumonia and finally localize the COVID-19 infected areas in the X-ray. This method was also based on a pre-trained VGG-16 model that was trained on the ImageNet dataset. Apostolopoulos and Mpesiana [22] evaluated different CNN architectures and adopted transfer learning to detect the abnormalities of chest X-rays and classify them to different classes. Authors demonstrated that VGG19 and MobileNet v2 are among the best CNN-based architectures to detect COVID-19 from X-ray images.

On the other hand, deep ResNet-based models have demonstrated an outstanding performance in COVID-19 detection. For instance, Oh et al. [39] proposed a patch-based CNN architecture based on simple ResNet-18 to train a limited size dataset for detecting COVID-19 from X-ray images and adopted the transfer learning using the pre-trained weights from ImageNet dataset. Similarly, Jain et al. [36] adopted transfer learning from ImageNet dataset and proposed a two-stage strategy using residual network architectures, i.e. ResNet50 and ResNet101. This method was trained and validated on medium size dataset and also utilized data augmentation. Yoo et al. [46] used deep learning-based decision-tree classifier to detect COVID-19 and tuberculosis from X-ray images. Specifically, they adopted a pre-trained ResNet18 on the ImageNet dataset and 3 binary decision trees for classification.

Several studies evaluated different CNN-based models to choose the most efficient one [25,38,47,54]. For example, Narin et al. [38] used 5 pre-trained CNNs-based architectures (i.e. ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2) and implemented 3 binary classifications to classify different lung diseases from X-ray images. Authors noted that, the pre-trained ResNet50 network was able to achieve the best classification performance compared with other CNNs models. Minaee et al. [25] proposed a deep learning architecture for COVID-19 detection based on fine-tuning of four state-of-the-art CNN models (ResNet18, ResNet50, SqueezeNet, and DenseNet121). Authors collected a dataset of 5000 chest X-rays to train and test their model and achieved good sensitivity and specificity rates. Also, Zebin and Rezvy [47] implemented a transfer learning pipeline and used 3 pre-trained CNN-based models (VGG16, ResNet50, and EfficientNetB0) for better feature extraction. This method was used to classify COVID-19 chest X-ray images collected from 2 publicly available datasets.

Recurrent neural networks (RNN) and LSTM were also widely adopted in literature for time series prediction of COVID-19. For instance, Demir [55] proposed a deep LSTM model to automatically identify COVID-19, pneumonia, and normal subjects from X-ray images. Islam et al. [35] adopted a CNN and an LSTM architecture to automatically diagnose COVID-19 from X-ray images. Specifically, CNN was adopted for deep features extraction while LSTM was used to detect COVID-19 subjects based on the extracted CNN features. Similarly, Khan et al. [56], proposed a hybrid framework based on CNN to extract multi time-scale features from convolutional layers and LSTM that identifies short, medium, and long-term dependencies by learning the representation of time-series data. Likewise, Aslan et al. [57] combined CNN with bidirectional LSTM model to improve the feature learning and boost the detection accuracy of COVID-19 from X-ray images. In another work, Aradhya et al. [58] proposed a cluster-based one-shot learning that requires few samples. Their model was based on an ensemble of the generalized regression neural network (GRNN) and the probabilistic neural network (PNN) classifiers at decision level.

To summarize, in Table 1 we present the most related works on COVID-19 diagnosis indicating the used architecture, dataset size, task(s), and the performance evaluation.

3. Material and methods

3.1. Multi-center datasets

In this work, we collected a total number of 11,312 X-ray radiographs from 10 multi-center datasets. In our dataset, we have 5656 and 5656 X-ray images for COVID-19 infected and healthy cases, respectively. The detailed information of multi-center datasets is given in Table 2. In our collected dataset, we randomly picked part of each dataset since some datasets have too many samples (e.g., D2 has 4235 samples and D5 has 4044 samples). In addition, we chose the same number of COVID-19 and healthy images to avoid the class imbalance problem. Examples of healthy and COVID-19 X-ray images from different datasets are shown in Fig. 1.

3.2. Proposed method

Our proposed framework consists of three main stages. First, we adopt the VGG architecture [20] to extract high level features from the X-ray radiographs. Then, we compute the mean values of the extracted features for the infected and healthy cases, respectively. Second, we use two computed mean values of extracted features as a supervisory factor to constrain the VGG network. In this part, we train the VGG network again to output the extracted features and an initial predicted score for every X-ray image. Third, by designing our multi-center COVID-19 graph in GCN, we accomplish the COVID-19 diagnose task. The overview of whole COVID-19 diagnosis framework is shown in Fig. 2. Note that, the words “multi-center” and “multi-site” are used interchangeably throughout this paper.

Table 1
Related works on COVID-19 diagnosis from chest X-ray radiographs.

Author(s)	Architecture	Dataset	Task	Performance
Zhang et al. [48]	CNN	100 COVID-19 1431 Other	COVID-19 vs. Others	96.0% (Sens) 70.65% (Spec) 0.952 (AUC)
Ozturk et al. [40]	CNN	127 COVID-19 500 No-findings 500 Pneumonia	COVID vs. No-Findings COVID vs. No-Findings vs. Pneumonia	98.08%(ACC) 87.02% (ACC)
Islam et al. [35]	CNN+LSTM	1525 COVID-19 3050 Other	COVID-19 vs. Normal vs. Pneumonia	99.4% (ACC) 99.9% (AUC) 99.2% (Spec) 99.3% (Sens) 98.9% (F1-score)
Chowdhury et al. [30]	CNN	219 COVID-19 1341 Normal 1345 viral pneumonia	COVID-19 vs. normal, vs. viral pneumonia	96.58% (ACC) 96.58% (Prec) 96.59% (recall) 96.58% (F1 score)
Panwar et al. [41]	VGG16 + CNN (nCOVnet)	142 COVID-19 142 Normal	COVID-19 vs. Normal	97.62% (Sens) 78.57% (Spec) 88.10% (ACC) 0.881 (AUC)
Apostolopoulos and Mpesiana [22]	Different CNNs	448 COVID-19 1414 bacterial pneumonia 1008 Normal	COVID-19 vs. Others	96.78% (ACC) 98.66% (Sens) 96.46% (Spec)
Oh et al. [39]	(FC)-DenseNet103 ResNet-18	191 Normal 57 tuberculosis 54 bacterial pneumonia 20 viral pneumonia, 180 COVID-19.	COVID-19 vs. Others	70.7 (ACC) 60.6 (Prec) 60.1 (Recall) 59.3 (F1 score) 89.7 (Spec)
Jain et al. [36]	ResNet50 and ResNet101	250 COVID-19 315 Normal 300 bacterial pneumonia 350 viral pneumonia.	COVID-19 vs. Others	98.93 (ACC) 98.93 (Sens) 98.66 (Spec) 96.39 (Prec) 98.15 (F1 score)
Narin et al. [38]	ResNet50, 101, ResNet152, InceptionV3 and Inception-ResNetV2	1023 COVID-19 2772 bacterial pneumonia 2800 Normal 1493 Viral Pneumonia	COVID-19, normal viral pneumonia and bacterial pneumonia	ResNet50 96.1% (ACC) Dataset-1, 99.5% (ACC) for Dataset2 and 99.7% (ACC) Dataset-3
Minaee et al. [25]	ResNet18, 50, SqueezeNet, DenseNet161	520 COVID-19 5000 Non COVID-19	COVID-19 vs. Others	98% (Sens) 90% (Spec)
Zebin and Rezvy [47]	VGG16, ResNet50, and EfficientNetB0	673 X-ray, CT 69% COVID-19 300 for others	COVID-19 vs. Normal	VGG16 90% (ACC) ResNet50 94.3% (ACC) EfficientNetB0 96.8% (ACC)
Waheed et al. [52]	CNN+GAN	403 COVID-19 721 Normal	COVID-19 vs. Normal	95% (ACC) 90% (Sens) 97% (Spec)
Abbas et al. [53]	CNN (ResNet)	105 COVID-19 80 Normal 11 SARS cases	COVID-19 vs. Normal vs. SARS	95.12 (ACC) 97.91 (Sens) 91.87 (Spec) 96.55 (AUC)
Aradhya et al. [58]	GRNN+PNN	69 COVID-19 79 Normal 79 bacterial pneumonia 79 viral pneumonia.	COVID-19 vs. Normal COVID-19 vs. Normal vs. bacterial pneumonia COVID-19 vs. Normal vs. bacterial pneumonia vs. viral pneumonia	100% (ACC) 85.23% (ACC) 74.05% (ACC)
Sethy et al. [59]	ResNet50+SVM	127 COVID-19 127 Pneumonia 127 Healthy	COVID-19 vs. Others	95.33% (ACC) 95.33% (Sens) 0.953 (F1 score)
Moutounet-Cartan [60]	VGG-16	125 COVID-19 152 Normal 48 Other pneumonia	COVID-19 vs. No Finding vs. Other Pneumonia	93.9% (ACC) 87.7% (COVID-19 Sens) 96.8% (No Finding Sens)

ACC=Accuracy, Sens=Sensitivity, Spec=Specificity, Prec=Precision, AUC=Area Under Curve.

3.3. Supervision mechanism

As shown in Fig. 2, our VGG network includes seven convolutional layers followed by seven max-pooling layers and rectified linear unit (ReLU) as an activation function. The detailed parameters of the convolutional layers and pooling layers are shown in Table 3. We have total of 11,312 images from 10 datasets, whereas there are a total 1,498,372 parameters in our network. Since our collected dataset comes from multi-center datasets, there is an indispensable difference among the X-ray radiographs. Hence, the parameters of VGG network result in noises on the extracted features. Significant fluctuations in the extracted features also deteriorate the performance of the classifier. To improve the quality of the extracted features and alleviate the effect of

Table 2
Detailed information of the multi-center datasets from which we collected our dataset.

Datasets	COVID-19	Healthy	Links
D1	700	700	https://github.com/ieee8023/covid-chestxray-dataset
D2	1500	1500	https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
D3	90	90	https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets
D4	900	900	https://data.mendeley.com/datasets/2fxz4px6d8/4
D5	1500	1500	https://data.mendeley.com/datasets/8h65ywd2jr/3
D6	200	200	https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
D7	56	56	https://www.kaggle.com/darshan1504/covid19-detection-xray-dataset
D8	130	130	https://www.kaggle.com/tarandeep97/covid19-normal-posteroanteriorpa-xrays
D9	30	30	https://www.kaggle.com/anaselmasry/covid19xray
D10	550	550	https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia
Total	5656	5656	

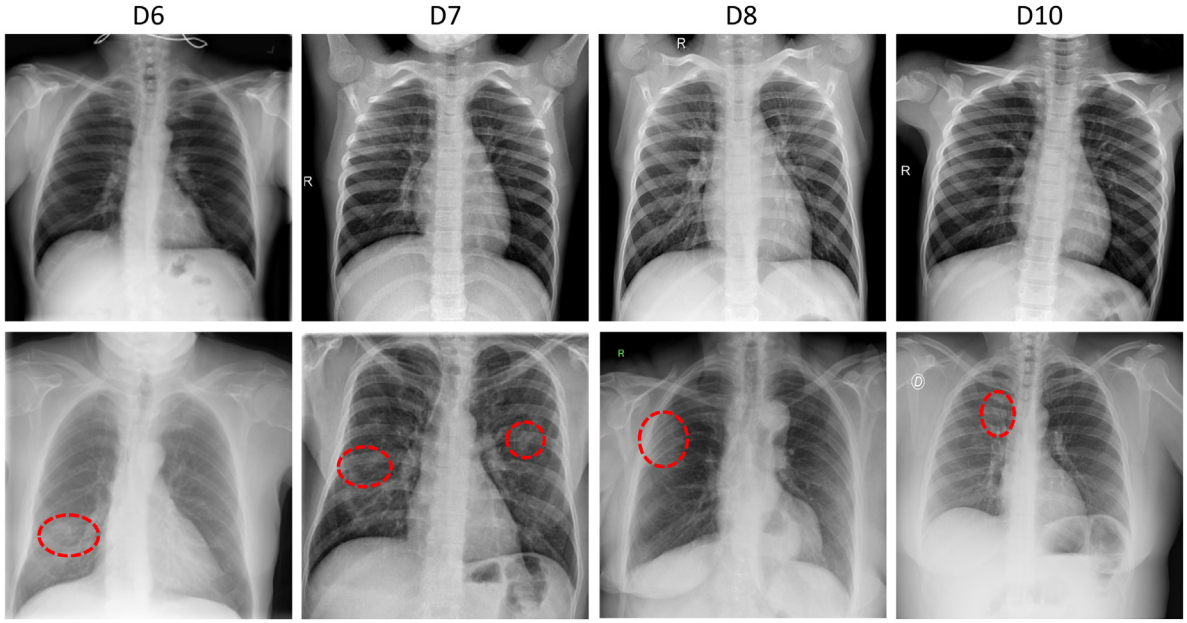


Fig. 1. Examples of healthy and COVID-19 X-ray images from different datasets. Columns from left to right represent D6, D7, D8, and D10, respectively. Rows from top to down represent normal and COVID-19 radiographs, respectively. (Some infected regions marked by chest physician are highlighted with red dashed circles in the second row).

the noise, we propose a supervision mechanism. The feature extraction process details are as follows: First, we pre-train the VGG network and use it to extract initial features. Specifically, we split the dataset into 80% and 20% for training and testing sets, respectively. In addition, during the training process, a 10% of the training samples are used for validation. Then, we use all training samples to pre-train our VGG network, and then use the pretrained VGG network to extract the features from the last pooling layer for every sample. Therefore, we get a 256×1 feature vector for every image. By averaging the feature vectors for COVID-19 cases and for healthy cases, respectively, we get a mean COVID-19 feature vector and a mean healthy feature vector. Second, we revise the VGG network by adding the supervision mechanism, which uses the mean COVID-19 feature vector and the mean healthy feature vector to constrain the extracted features to realize the aim of suppressing noises. In this part, we train the proposed VGG framework with the supervision mechanism again. The training and testing samples are same with that in the first step. By using the trained VGG framework, we finally get the updated extracted features for every image and get an initial score for every image to represent the initial diagnosis results.

For training samples, we compute the mean values of feature vectors and get a mean feature vector for COVID-19 infected cases and a mean feature vector for healthy patients. Using the above two mean feature vectors, we design a supervision mechanism on the VGG network as shown in Fig. 2.

3.4. VGG network for feature extraction and initial scores

As shown in Fig. 2, our VGG network with the supervision mechanism considers the feature's difference of the COVID-19 and healthy cases. Since the training samples' status are already known, we use the two kinds of the mean feature values to constrain the extracted features. Specifically, we use them to extract supervised features and get an initial score for every X-ray radiograph by training it again. To further reduce the dimension of the feature vector, we use the recursive feature elimination [61] to select the most discriminative features represented in a low-dimensional feature vector for every subject.

3.5. Graph convolutional network

Compared with traditional neural networks, GCN uses graph theory to improve performance, particularly it plays a quite notable role in filtering noise. Typically, each node in the graph represents the feature vector of a subject and the edge represents the interaction between corresponding pair of nodes on a given graph. Graph theory uses all nodes on the graph to perform convolution, and edge weights are the key to its performance as

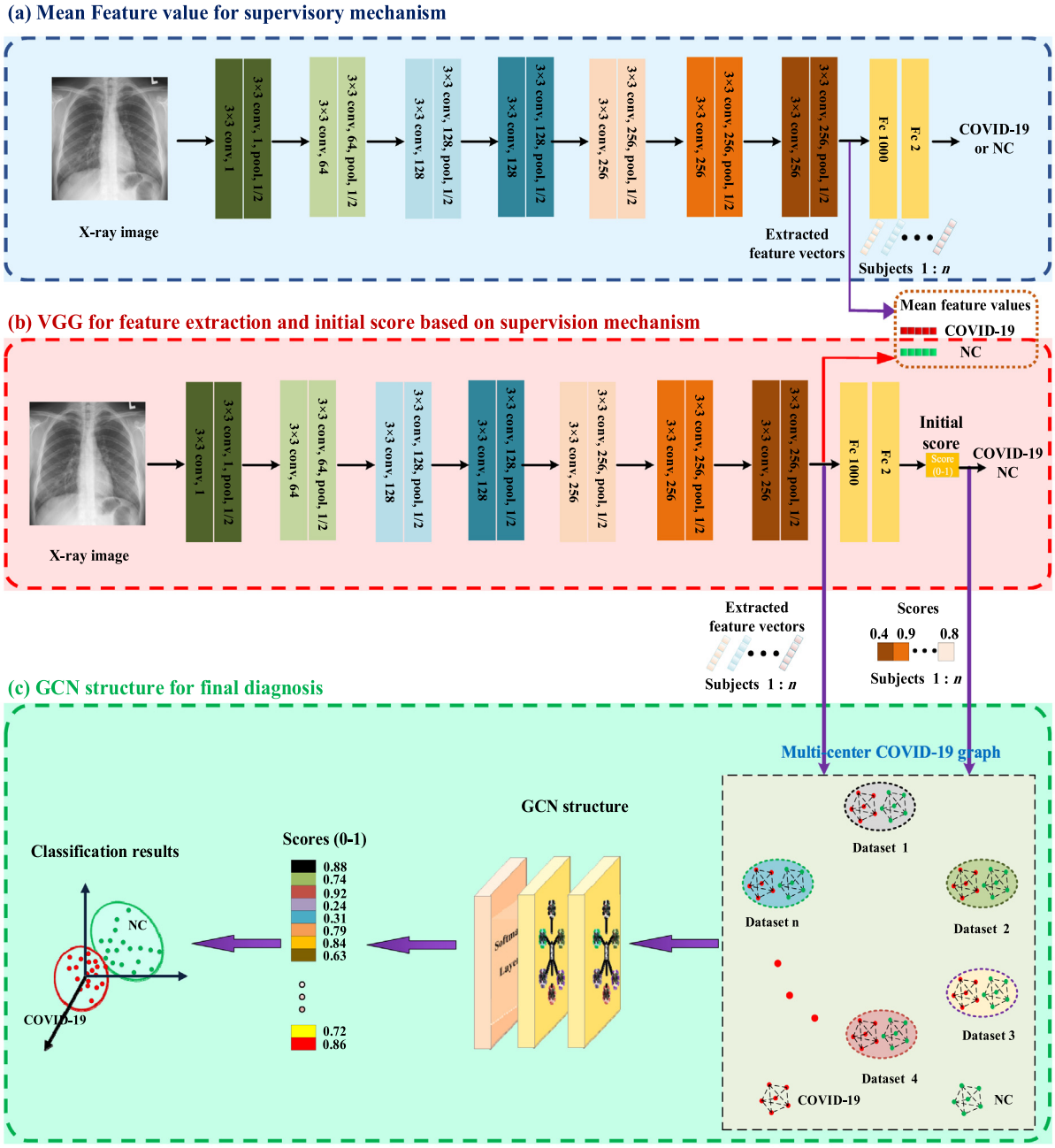


Fig. 2. Overview of the proposed COVID-19 diagnosis framework. (a) The VGG network to extract initial features from X-ray images to compute mean feature values. (b) The supervision mechanism based on the mean feature values on VGG network to improve its performance and use it to extract the feature vector and the initial score for every X-ray image. (c) The multi-center COVID-19 GCN that considers dataset information, status of training samples, and initial scores. Softmax layer of the GCN is used to score every X-ray radiograph and output the final diagnosis results.

they are corresponding to convolution coefficients. Thereby, they attracted much attention [62–64]. In the following, we propose a multi-center COVID-19 graph in GCN to establish edges to fit the diagnosis task's characteristics.

3.5.1. Multi-center COVID-19 graph

We first examine the differences among the datasets; hence we split all nodes on the proposed GCN into several groups (clusters). In other words, subjects in the same dataset represent a cluster, which means the number of clusters represent the number of datasets. In this paper, total of 10 clusters are corresponding to the 10 datasets. For the nodes in different group, we do not set up connections for them that means their edges will have zero weights. Regarding the nodes within the same group, we set the weights of their edges by considering the training samples' disease status and initial scores from the VGG network. Specifically, the edge weight between two subjects in the same group is calculated using the following formula:

$$\boldsymbol{A}(v,u) = (a * Sim(\boldsymbol{F}_v, \boldsymbol{F}_u) + b * Sim(\boldsymbol{Score}_v, \boldsymbol{Score}_u)) \times (1 + r_S(S_v, S_u)), \quad (1)$$

$$r_S(S_v, S_u) = \begin{cases} 1, & S_v = S_u \\ 0, & otherwise, \end{cases} \quad (2)$$

where the edge weights construct the adjacency matrix ($\boldsymbol{A}$). $\boldsymbol{A}(v,u)$ represents the edge weight between subjects v and u,

Table 3
Detailed parameters of the proposed network.

Layer	Type	Filter size	Stride	Kernel	Input size	No. of parameters
1	Convolution2D	3 × 3	1	64	128 × 128	150
2	Convolution2D	3 × 3	1	128	128 × 128	2040
3	Pool	2 × 2	2	128	64 × 64	0
4	Convolution2D	3 × 3	1	256	64 × 64	3400
5	Convolution2D	3 × 3	1	256	64 × 64	5650
6	Pool	2 × 2	2	256	32 × 32	0
7	Convolution2D	3 × 3	1	64	32 × 32	11 300
9	Convolution2D	3 × 3	1	128	32 × 32	22 550
10	Pool	2 × 2	2	128	16 × 16	0
11	Convolution2D	3 × 3	1	256	16 × 16	22 550
12	Convolution2D	3 × 3	1	256	16 × 16	22 550
13	Pool	2 × 2	2	256	8 × 8	0
.14	Convolution2D	3 × 3	1	64	8 × 8	45 100
15	Convolution2D	3 × 3	1	128	8 × 8	90 100
16	Pool	2 × 2	2	128	4 × 4	0
17	Convolution2D	3 × 3	1	256	4 × 4	18 020
18	Convolution2D	3 × 3	1	256	4 × 4	360 200
19	Pool	2 × 2	2	256	2 × 2	0
20	Convolution2D	3 × 3	1	64	2 × 2	360 200
21	Convolution2D	3 × 3	1	128	2 × 2	360 200
22	Pool	2 × 2	2	128	1 × 1	0
23	Fc1					12 060
24	Fc2					122
						1,498,372

$Sim(\cdot)$ is the similarity of imaging information, $\boldsymbol{F}_v$ and $\boldsymbol{F}_u$ are the extracted feature vectors of subjects v and u, respectively. a and b are weight parameters of the similarity terms. $\boldsymbol{Score}_v$ and $\boldsymbol{Score}_u$ denote the scores of subjects v and u from the VGG network, r_S represents the distance of the disease status (the statuses of those training samples on graph are known) while S_v and S_u represent their disease status.

The most widely used method for evaluating similarity of imaging information based on the extracted feature is based on the correlation distance which is calculated using [65]:

$$Sim\left(\boldsymbol{F}_v, \boldsymbol{F}_u\right) = exp\left(-\frac{\left[\rho\left(\boldsymbol{F}_v, \boldsymbol{F}_u\right)\right]^2}{2\sigma^2}\right), \tag{3}$$

where $\rho(\cdot)$ and σ are the correlation distance function and the width of the kernel, respectively.

Calculating the correlation distance using Eq. (3) may affect the convolution performance. Therefore, we estimate the similarity measure based on the fact that, the VGG network has good feature extraction capability. By using VGG network to extract features from images, we get the diagnosis score of each subject, see Fig. 2. Based on the calculated scores, the similarities can be computed using:

$$Sim\left(\boldsymbol{Score}_v, \boldsymbol{Score}_u\right) = exp\left(-\frac{\left[\rho\left(\boldsymbol{Score}_v - \boldsymbol{Score}_u\right)\right]^2}{2\sigma^2}\right). \tag{4}$$

3.5.2. GCN structure

In the GCN-based methods, the spectral theory applies the convolution of Fourier transform and Taylor's expansion formula to process the adjacency matrix to achieve a better filtering effect and computational efficiency [65]. To calculate the spectral convolution, we can consider it as a multiplication operation of a signal $\boldsymbol{x} \in \mathbb{R}^n$ and a filter $g_{\boldsymbol{\theta}} = diag(\boldsymbol{\theta})$ as follows:

$$g_{\boldsymbol{\theta}} * \boldsymbol{x} = \boldsymbol{U} g_{\boldsymbol{\theta}}\left(\boldsymbol{\Lambda}\right) \boldsymbol{U}^T \boldsymbol{x} = \sum_{k=0}^{K} \boldsymbol{\theta}_k T_k\left(\tilde{\boldsymbol{L}}\right) \boldsymbol{x}, \tag{5}$$

where $\boldsymbol{U}$ is the matrix of eigenvectors, $\boldsymbol{\theta}_k$ is a vector of Chebyshev coefficients, T_k is Chebyshev polynomials function, $\tilde{\boldsymbol{L}} = 2/\lambda_{max}\boldsymbol{\Lambda} - \boldsymbol{I}_N$, and $g_\theta(\boldsymbol{\Lambda})$ is well approximated by a truncated expansion in terms of Chebyshev polynomials to the K^{th}-order. In this paper, we use the following formula to calculate $\boldsymbol{U}$:

$$\boldsymbol{L} = \boldsymbol{I}_N - \boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{A}_{ac} \boldsymbol{D}^{-\frac{1}{2}} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^T, \tag{6}$$

where $\boldsymbol{I}_N$ represents the identity matrix whilst $\boldsymbol{D}$ is the diagonal degree matrix.

After performing the spectral convolution operation, the adjacency matrix $\boldsymbol{A}_s$ is estimated using $\sum_{k=0}^{K} \boldsymbol{\theta}_k T_k(\tilde{\boldsymbol{L}})$ where K is the polynomial order and it is used to control the filter effect (see its effect in Section 5.2).

Finally, the proposed architecture combines two graph convolutional layers that use the ReLU as an activation function (See Fig. 2(c)). Following the two graph convolutional layers, a Softmax function is used to output probability distribution for the final diagnosis of the disease.

4. Experiments and results

4.1. Experimental setup and evaluation metrics

All experiments of this paper were conducted on a machine with Intel(R) CPU i7-8700 at 3.19 GHz, GPU NVIDIA TITAN Xp, 128G of RAM, using Keras deep learning library. In our VGG network, we used the following parameters. Learning rate was set to $1e^{-4}$, Adam optimizer was adopted for optimization, and the number of epochs was 200. In addition, our dataset was divided to 80% for training (in which 10% were used for validation) and the remaining 20% for testing. The total running time is composed of the training of the three phases of our framework; the VGG feature extraction, the supervision mechanism, the GCN diagnosis (Fig. 2(a)–(c)). The training time for these phases was, respectively, 2.27 h, 1.84 h, and 0.4 h (totally 4.11 h).

In this work, we evaluated the performance of our methods using several metrics. Namely, we measured the sensitivity (Sens), specificity (Spec), accuracy (ACC), and area under the curve (AUC) of receiver operating characteristic (ROC) which are, respectively, defined as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}, \tag{7}$$

$$Sens = \frac{TP}{TP + FN}, \tag{8}$$

$$Spec = \frac{TN}{FP + TN}, \tag{9}$$

where TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively.

4.2. Comparison methods

In this paper, we used several deep learning-based methods to perform feature extraction. These methods are based on CNN, ResNet, and VGG architectures. Additionally, we used these methods with and without the proposed supervision mechanism to evaluate its efficacy. For classification, we also considered different classifiers. Namely, multilayer perceptron (MLP) [66], random forests (RF) [67], gradient boosted decision trees (GBDT) [68], GCN, and the proposed GCN. In addition, we adopted the end-to-end training as baseline to evaluate the effect of the proposed supervision mechanism on feature learning. The reason of choosing these methods is that, they have been widely used in literature and achieved the state-of-the-art performance in many tasks. For GCN and MLP, the main parameters that were used in the comparisons are as follows. Number of iterations was 3, dropout rate was 0.1, l_2 regularization was 5×10^{-4}, learning rate was 0.005, number of epochs was 200, number of neurons

per layer was 32, and number of extracted features was 30. For RF parameters, the number of trees was 500, and the maximum depth was three. For GBDT, learning rate was 0.01, number of epochs was 200, and maximum depth of the tree was 5.

4.3. Classification performance

In Table 4, we summarize the classification results for our binary task (COVID-19 vs. Healthy) using different feature extraction methods (with and without the supervision mechanism) and classifiers. It can be noticed that, the proposed supervision mechanism effectively improves the performance of all classifiers. In particular, the proposed GCN with the supervision mechanism on the VGG network achieves 96.41%, 96.6%, 96.2%, and 98.7% for ACC, Sens, Spec, and AUC, respectively. Best results are given in boldfaces for every classifier. In addition, we compared the performance of our method with the state-of-the-art methods for COVID-19 diagnosis. Comparisons are given in Table 5. From this Table, it is obvious that the proposed method is not only trained on the biggest dataset but also it achieves the best performance in COVID-19 diagnosis. Note that, although the dataset used in Wang et al. [44] is apparently the biggest dataset, the major part of this dataset is generated one, not real X-ray images. Similarly, Luz et al. [51] used a quite big dataset of 13,569 subjects (but only from 3 centers), yet, our performance has about 2.5% higher accuracy than it. Despite our proposed method achieves higher accuracy than the methods in Table 5, there are other methods in literature that attained better accuracy than ours e.g., [35,36,40,54,55,57,58]. Nevertheless, these methods were tested on a relatively small datasets with limited number of centers which may not be appropriate and thus hinder their clinical significance, without further investigations. It is noteworthy that, in our dataset, we maintain the class balance by choosing equal number of the healthy and the COVID-19 subjects (see Table 2) which not only neutralizes the feature learning process but also further promotes the reliability of the proposed method performance. Note that, the results in Table 4 were obtained by setting a and b in Eq. (1) to 0.5.

In addition, in Fig. 3, we depict the performance of the proposed method for the both accuracy and the loss function during the training phase. It is shown that, the training and validation accuracies of our method reach 99.01% and 96.41%, respectively, at epoch 64 (Fig. 3a). Similarly, our loss function (categorical crossentropy) has a stable behavior with almost same generalization gap around epoch 98 (Fig. 3b). The stability of the accuracy and the loss function during training process is mainly due to the effectiveness of our proposed supervision mechanism which enhances the feature learning ability in our model.

We further draw the ROC curves to demonstrate the effectiveness of the proposed supervision mechanism on the performance of different classifiers as shown in Fig. 4. It is clear that the classification performance is improved using our supervision mechanism, particularly with the proposed GCN.

5. Discussion

In this Section, we discuss the main advantages of proposed method and the current limitations in our study. Besides, we also investigate the effect of the supervision mechanism on the feature extraction, the polynomial order, and the number of features as well.

5.1. Effect of the supervision mechanism

The proposed supervision mechanism obviously improved the feature extraction procedure, not only the proposed method, but also many other deep feature extraction methods. This is demonstrated on a variety of performance metrics in Table 4. In this table, a number of feature extraction methods (2nd column) with and without the supervision mechanism are evaluated using various classifiers (1st column). It is shown that the supervision mechanism can boost the performance of all these classifiers, especially in terms of accuracy. For instance, in the last block of Table 4, the performance of our proposed GCN classifier with CNN feature extraction can be improved by 0.7%, 2.4%, and 1.9% in terms of ACC, Sens, and AUC, respectively, when the supervision mechanism is adopted.

We also demonstrate how the top 12 features are affected when using the proposed supervision mechanism in Fig. 5. In this figure, features of COVID-19 (yellow and red) and healthy (blue and green) subjects with and without supervision are plotted. It can be noticed that, there is a great difference between COVID-19 and healthy subjects when supervision mechanism is adopted as compared with no supervision. This can ultimately ease and boost the classification task by using the underlying classifier.

Moreover, we also evaluate the effect of weightings of the supervision mechanism on the accuracy of COVID-19 diagnosis. Table 6 lists the accuracy over different weighting values. In this experiment, the best accuracy was achieved at a supervision weighting of 0.2, with an increase of 1.81%.

5.2. Effect of the polynomial order

The polynomial order (K) in Eq. (5) also affects the performance and shall be carefully determined. In this subsection, we explore the effect of this parameter on the accuracy measure (ACC). Specifically, we empirically tested different values of K such that $K\epsilon\{1, 2, 3, 4, 5\}$. Table 7 reports the accuracy of different orders of K. From this table, it is noticed that, the accuracy slightly increases at higher degrees till it reaches the best performance at $K = 3$. However, choosing higher order of K ($K > 3$) decreases the diagnosis accuracy. Therefore, in this work, we fixed the polynomial order to 3 as it achieved the best performance.

5.3. Effect of number of features

We also explored the effect of the number of extracted features on the diagnosis performance. We tested the influence of different number of the extracted features N on the classification accuracy. In this experiment, we tested $N\epsilon[10, 250]$ and the classification accuracy is given in Table 8. It is obvious that, when $N \geq 30$, the accuracy is better than $N < 30$. On the other hand, the classification accuracy becomes the best (96.41%) when $N = 30$. By increasing the number of features, the accuracy slightly decreases. Hence, we fixed N to 30 features.

5.4. Limitations

Although the proposed GCN and the supervision mechanism achieved promising performance, the current study still has few limitations. First, we used only X-ray radiographs to extract COVID-19 biomarkers for disease diagnosis. However, using multimodal data fusion from CT and X-ray can potentially improve the diagnosis performance. Second, we only focused on binary classification task of COVID-19 and healthy subjects. Multiclass classification of other chest disease (e.g. pneumonia, viral pneumonia, and nodules) are increasingly desirable. Thus, it would be

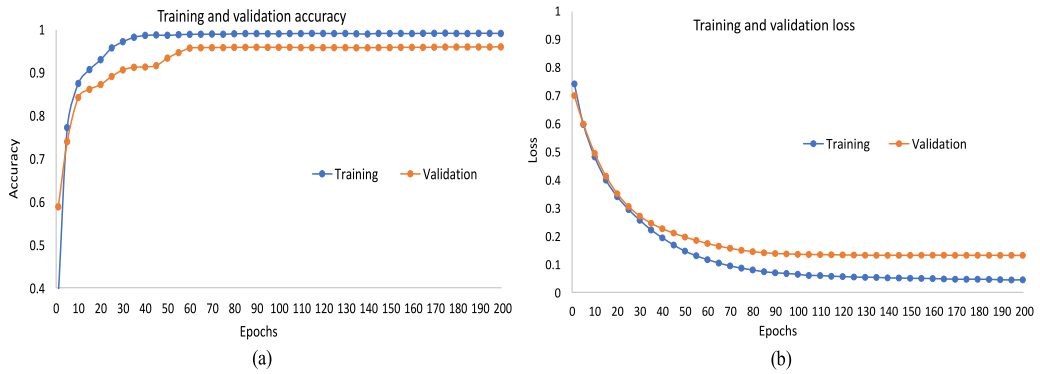


Fig. 3. Performance of the proposed method during training phase. (a) Accuracy (b) Loss function.

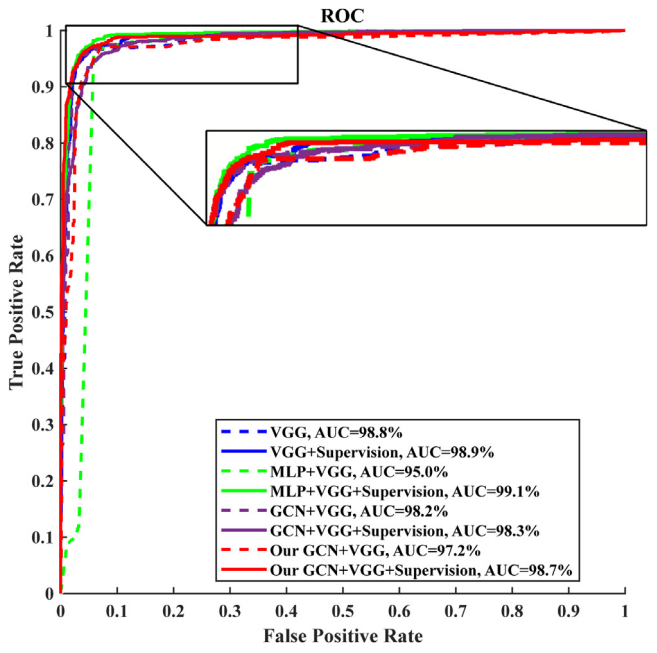


Fig. 4. ROC curves of the proposed GCN and the other classifiers using our supervision mechanism. The solid lines represent the classifiers performances using the supervision mechanism while the dashed lines represent the performances without the supervision mechanism. From the zoomed version on right, it is clear that supervision mechanism can enhance the diagnosis performance not only of our GCN but also for other methods as well.

Table 4
Classification performance using the proposed GCN and the supervision mechanism against different deep feature extraction and classification methods.

Classifier	Feature extraction	ACC (%)	Sens (%)	Spec (%)	AUC (%)
None	CNN	95.5	95.9	95.1	99.2
	CNN+Supervision	**96.2**	94.8	**97.5**	**99.4**
	Resnet	93.9	90.3	97.4	98.6
	Resnet+Supervision	94.3	91.9	96.8	98.7
	VGG	94.5	92.1	96.9	98.8
	VGG +Supervision	96.1	**96.4**	95.5	98.9
MLP	CNN	95.8	95.2	**96.3**	98.5
	CNN+Supervision	**96.1**	**96.8**	95.4	99.2
	Resnet	94.5	95.8	93.2	98.1
	Resnet+Supervision	95.0	96.2	93.8	98.4
	VGG	94.6	95.2	94.1	95.0
	VGG+Supervision	96.0	96.4	95.5	**99.1**
RF	CNN	93.3	94.7	91.8	92.8
	CNN+Supervision	92.2	**98.8**	85.6	91.6
	Resnet	83.2	85.4	80.8	82.7
	Resnet+Supervision	93.4	95.1	91.7	93
	VGG	84.3	84.3	84.4	83.9
	VGG+Supervision	**95.9**	96.2	**95.5**	**95.4**
GBDT	CNN	93.1	95.1	91.3	93.3
	CNN+Supervision	92.2	**98.9**	85.5	91.5
	Resnet	83.5	84.9	82.0	83.0
	Resnet+Supervision	93.4	95.1	91.6	92.9
	VGG	83.5	84.9	82.0	83.0
	VGG+Supervision	**95.6**	96.0	**95.3**	**95.2**
GCN	CNN	95.5	94.7	96.3	98.1
	CNN+Supervision	**96.2**	96.9	95.5	**99.0**
	Resnet	94.2	94.2	94.2	98.2
	Resnet+Supervision	94.3	**97.1**	91.5	98.0
	VGG	94.6	95.6	93.7	98.2
	VGG+Supervision	95.9	96.2	**95.7**	98.3
Our GCN	CNN	95.7	94.5	97	96.9
	CNN+Supervision	96.4	**96.9**	95.9	**98.8**
	Resnet	94.4	94.4	94.3	98.5
	Resnet+Supervision	94.8	93.0	**96.6**	98.5
	VGG	94.6	95.3	93.9	97.2
	VGG+Supervision	**96.41**	96.6	96.2	98.7

better to include other disease classes in our proposed method. Third, we only focused on the classification without quantifying the disease status. Indeed, it is quite advantageous to estimate the severity level of the infection towards more personalized treatment.

6. Conclusions and perspectives

In this paper, we proposed a novel method based on GCN with a supervision mechanism to diagnose COVID-19 from X-ray radiographs. Specifically, we combined the supervision mechanism with VGG network to extract the most informative features and used the proposed GCN to perform the classification to either COVID-19 or healthy images. Our method was trained on a relatively big dataset collected from a large number of sites/centers and attained superior performance to many other classifiers. This is particularly due to the effective feature learning capability of the proposed supervision mechanism. Moreover, the proposed method outperformed the state-of-the-art methods and achieved 96.41 %, 96.6%, 96.2%, and 98.7% for ACC, Sens, Spec, and AUC, respectively.

The eminent performance of our proposed method can help in combating the COVID-19 pandemic by providing an early diagnosis tool from chest X-ray radiographs, particularly in the low-income and isolated regions where RT-PCR test kits are not available. After further investigations, we believe that the proposed method can be adopted in the healthcare systems in different ways, especially with the development of the internet of things. One way to achieve this is incorporating our method with the screening system of the chest X-ray scanner to provide the radiologists/technicians with preliminary diagnosis. Another

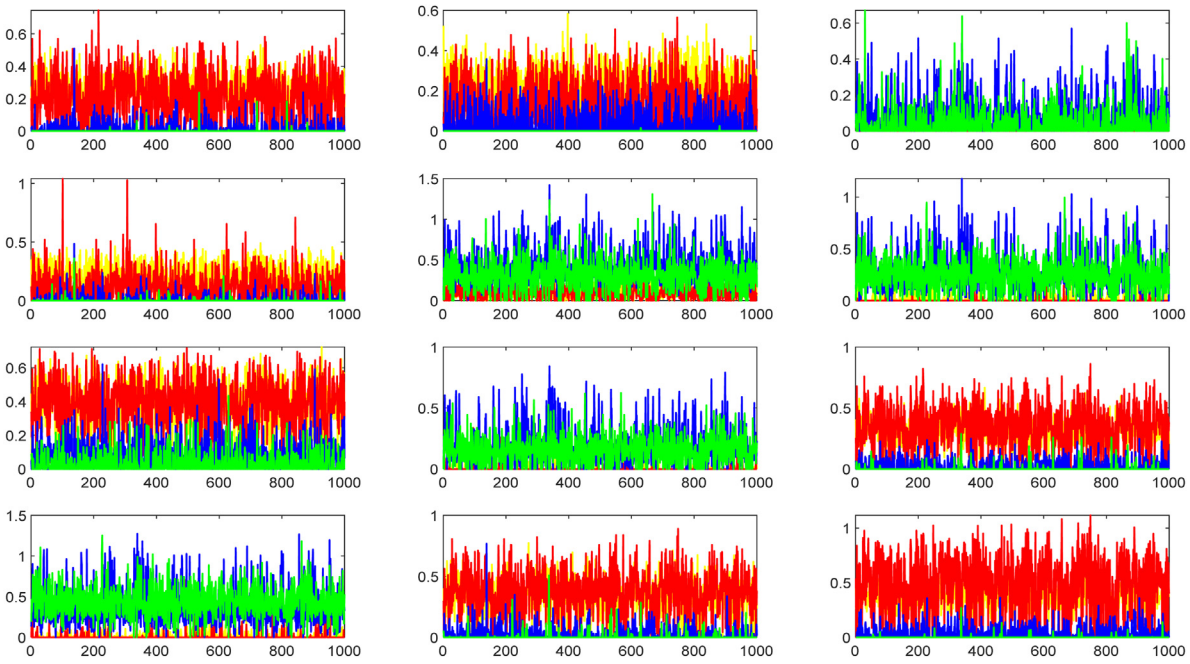


Fig. 5. Effect of the supervision mechanism on the top 12 features. Yellow and red colors are COVID-19 features while the blue and green colors represent the healthy features with and without supervision, respectively.

Table 5
Comparisons between the proposed method and the state-of-the-art COVID-19 diagnosis methods from X-ray radiographs.

Method	Dataset size	No. of centers	ACC (%)	Sens (%)	Spec (%)	AUC (%)
Zebin and Rezvy [47]	673	2	94.3	–	–	–
Panwar et al. [41]	284	1	88.1	97.62	78.57	88.09
Wang et al. [44]	13,975	5	93.3	91.0	–	–
Sethy et al. [59]	381	2	95.33	95.33	–	–
Waheed et al. [52]	1124	3	95	90	**97**	–
Abbas et al. [53]	196	2	95.12	**97.91**	91.87	96.55
Luz et al. [51]	13,569	3	93.9	96.8	–	–
Moutounet-Cartan [60]	327	2	93.9	87.7	–	97.4
Proposed	11,312	10	**96.41**	96.6	96.2	**98.7**

Table 6
Effect of weightings of the supervision on the COVID-19 diagnosis accuracy.

Predict/supervision	ACC (%)
1.0–0.0	94.6
0.9–0.1	96.0
0.8–0.2	**96.41**
0.7–0.3	96.1
0.6–0.4	94.8
0.5–0.5	95.2
0.4–0.6	95.6
0.3–0.7	96.1
0.2–0.8	96.0
0.1–0.9	93.8

Table 7
Effect of different degrees of the polynomial (K) on the diagnosis accuracy of COVID-19.

K	ACC (%)
1	96.36
2	96.36
3	**96.41**
4	96.06
5	96.10

Table 8
Effect of number of features on the diagnosis accuracy of proposed method.

Number of features (N)	ACC (%)
10	94.08
20	96.27
30	**96.41**
50	**96.41**
70	96.14
90	96.23
110	96.32
130	96.19
150	96.32
170	96.32
190	96.28
210	96.23
230	96.23
250	96.23

possible way is utilizing the proposed method (together with other patient's clinical data) in remote computer-aided diagnosis system on a cloud system platform. Finally, it might also be useful to develop a light-weight application of the proposed method to be applicable on hand-held devices.

In our future work, we will evaluate the performance of the proposed method on multimodal data (i.e. CT images and X-ray radiographs) towards learning complementary features. In addition, it will be interesting to measure the severity of the COVID-19 on the chest tissues to evaluate the risk of the infection. Last but not least, it will be highly desirable to learn the spread of the infection in the tissue. This can be achieved by modeling the volume changes of the infected areas and can be potentially addressed via adversarial learning.

CRediT authorship contribution statement

Ahmed Elazab: Conceptualization, Methodology, Data collection, Writing, Editing, Fund acquisition, Project management. **Mohamed Abd Elfattah:** Data collection, Writing, Editing, Organizing, Discussing. **Yuexin Zhang:** Conceptualization, Methodology, Data collection, Writing, Review.

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.

Acknowledgment

Authors would like to thank Prof. Chunqi Chang for his efforts in proofreading the manuscript and appreciate his constructive comments during the revision. Also, we would like to thank Dr. Xuegang Song for providing his source code and his assistance to get this work completed.

This work was supported by National Natural Science Foundation of China (Nos. 61950410615, 62101338, and 61971289), China Postdoctoral Science Foundation (No. 2019M653014), Shenzhen Fundamental Research Project (JCYJ20170412111316339), and Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions.

References

[1] https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19 (Last Access; October 11th, 2021), 2021.
[2] https://covid19.who.int/ (Last Access; October 11th, 2021), 2021.
[3] C.J.E. Metcalf, D.H. Morris, S.W. Park, Mathematical models to guide pandemic response, Science 369 (6502) (2020) 368–369.
[4] N. Zhu, et al., A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. (2020).
[5] S. Lee, et al., Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with SARS-CoV-2 infection in a community treatment center in the Republic of Korea, JAMA Intern. Med. (2020).
[6] L. Mousavizadeh, S. Ghasemi, Genotype and phenotype of COVID-19: Their roles in pathogenesis, J. Microbiol. Immunol. Infect. (2020) 2020/03/31.
[7] T. Ai, et al., Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases, Radiology (2020) 200642.
[8] L. Zhou, et al., A rapid, accurate and machine-agnostic segmentation and quantification method for CT-based COVID-19 diagnosis, IEEE Trans. Med. Imaging 39 (8) (2020) 2638–2652.
[9] S.A. Lauer, et al., The incubation period of coronavirus disease (COVID-19) from publicly reported confirmed cases: estimation and application, Ann. Intern. Med. 172 (9) (2020) 577–582.
[10] C. Wang, A. Elazab, J. Wu, Q. Hu, Lung nodule classification using deep feature fusion in chest radiography, Comput. Med. Imaging Graph. 57 (2017) 10–18.
[11] A. Zali, et al., Correlation between low-dose chest computed tomography and RT-PCR results for the diagnosis of COVID-19: A report of 27, 824 cases in tehran, Iran, Acad. Radiol. (2020) 2020/09/21.
[12] O.S. Albahri, et al., Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects, J. Infect. Public Health 13 (10) (2020) 1381–1396.
[13] S.K. Zhou, et al., A review of deep learning in medical imaging: Image traits, technology trends, case studies with progress highlights, and future promises, 2020, arXiv preprint arXiv:2008.09104.
[14] M. Jamshidi, et al., Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment, IEEE Access 8 (2020) 109581–109595.
[15] L. Wynants, et al., Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal, Bmj 369 (2020).
[16] S. Lalmuanawma, J. Hussain, L. Chhakchhuak, Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review, Chaos Solitons Fractals (2020) 110059.
[17] F. Shi, et al., Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19, IEEE Rev. Biomed. Eng. (2020).
[18] H. Chiroma, A.E. Ezugwu, F. Jauro, M.A. Al-Garadi, I.N. Abdullahi, L. Shuib, Early survey with bibliometric analysis on machine learning approaches in controlling COVID-19 outbreaks, Vol. 6, 2020, p. e313.
[19] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[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] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, arXiv preprint arXiv:1409.1556.
[22] I.D. Apostolopoulos, T.A. Mpesiana, Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks, Phys. Eng. Sci. Med. (2020) 1.
[23] P. Chhikara, P. Singh, P. Gupta, T. Bhatia, Deep convolutional neural network with transfer learning for detecting pneumonia on chest X-rays, in: Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals, Springer, 2020, pp. 155–168.
[24] T. Mahmud, M.A. Rahman, S.A. Fattah, CovXNet: A Multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization, Comput. Biol. Med. 122 (2020) 103869.
[25] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G. Jamalipour Soufi, Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning, Med. Image Anal. 65 (2020) 101794, 2020/10/01.
[26] J. Zhu, B. Shen, A. Abbasi, M. Hoshmand-Kochi, H. Li, T.Q. Duong, Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs, PLoS One 15 (7) (2020) e0236621.
[27] I.D. Apostolopoulos, S.I. Aznaouridis, M.A. Tzani, Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases, J. Med. Biol. Eng. (2020) 1.
[28] L. Brunese, F. Mercaldo, A. Reginelli, A. Santone, Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays, Comput. Methods Programs Biomed. 196 (2020) 105608.
[29] X. Chen, L. Yao, Y. Zhang, Residual attention U-net for automated multi-class segmentation of COVID-19 chest CT images, 2020, arXiv preprint arXiv:2004.05645.
[30] N.K. Chowdhury, M.M. Rahman, M.A. Kabir, PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images, Health Inf. Sci. Syst. 8 (1) (2020) 1–14.
[31] D. Dansana, et al., Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm, Soft Comput. (2020) 1–9.
[32] G. Gaál, B. Maga, A. Lukács, Attention u-net based adversarial architectures for chest x-ray lung segmentation, 2020, arXiv preprint arXiv:2003.10304.
[33] A. Haghanifar, M.M. Majdabadi, S. Ko, Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning, 2020, arXiv preprint arXiv:2006.13807.
[34] E.E.-D. Hemdan, M.A. Shouman, M.E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, 2020, arXiv preprint arXiv:2003.11055.
[35] M.Z. Islam, M.M. Islam, A. Asraf, A combined deep cnn-lstm network for the detection of novel coronavirus (covid-19) using X-ray images, Inform. Med. Unlocked 20 (2020) 100412.
[36] G. Jain, D. Mittal, D. Thakur, M.K. Mittal, A deep learning approach to detect Covid-19 coronavirus with X-ray images, Biocybern. Biomed. Eng. (2020).
[37] A.I. Khan, J.L. Shah, M.M. Bhat, Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, Comput. Methods Programs Biomed. (2020) 105581.
[38] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, 2020, arXiv preprint arXiv:2003.10849.
[39] Y. Oh, S. Park, J.C. Ye, Deep learning covid-19 features on cxr using limited training data sets, IEEE Trans. Med. Imaging (2020).
[40] 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.
[41] H. Panwar, P. Gupta, M.K. Siddiqui, R. Morales-Menendez, V. Singh, Application of deep learning for fast detection of COVID-19 in X-rays using ncovnet, Chaos Solitons Fractals (2020) 109944.
[42] S. Rajaraman, S. Antani, Weakly labeled data augmentation for deep learning: A study on COVID-19 detection in chest X-rays, Diagnostics 10 (6) (2020) 358.
[43] F. Ucar, D. Korkmaz, COVIDiagnosis-NEt: Deep Bayes-SqueezeNet based diagnostic of the coronavirus disease 2019 (COVID-19) from X-ray images, Med. Hypotheses (2020) 109761.

[44] L. Wang, A. Wong, COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images, 2020, arXiv preprint arXiv:2003.09871.
[45] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, J. Liu, Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing, Radiology (2020) 200343.
[46] S.H. Yoo others, Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging, Front. Med. 7 (2020) 427.
[47] T. Zebin, S. Rezvy, COVID-19 Detection and disease progression visualization: Deep learning on chest X-rays for classification and coarse localization, Appl. Intell. (2020) 1–12.
[48] J. Zhang, Y. Xie, Y. Li, C. Shen, Y. Xia, Covid-19 Screening on chest x-ray images using deep learning based anomaly detection, 2020, arXiv preprint arXiv:2003.12338.
[49] X. Li, C. Li, D. Zhu, COVID-MobileXpert: ON-device COVID-19 screening using snapshots of chest X-ray, 2020, arXiv preprint arXiv:2004.03042.
[50] H. Mukherjee, S. Ghosh, A. Dhar, S.M. Obaidullah, K.C. Santosh, K. Roy, Shallow convolutional neural network for COVID-19 outbreak screening using chest X-rays, Cogn. Comput. (2021) 2021/02/05.
[51] E. Luz, P.L. Silva, R. Silva, L. Silva, G. Moreira, D. Menotti, Towards an effective and efficient deep learning model for covid-19 patterns detection in x-ray images, 2020.
[52] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, P.R. Pinheiro, CovidGAN: Data augmentation using auxiliary classifier GAN for improved Covid-19 detection, IEEE Access 8 (2020) 91916–91923.
[53] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, Appl. Intell. 51 (2) (2021) 854–864, 2021/02/01.
[54] P. Kedia, Anjum, R. Katarya, CoVNet-19: A Deep learning model for the detection and analysis of COVID-19 patients, Appl. Soft Comput. 104 (2021) 107184, 2021/06/01.
[55] F. Demir, DeepCoroNet: A Deep LSTM approach for automated detection of COVID-19 cases from chest X-ray images, Appl. Soft Comput. 103 (2021) 107160, 2021/05/01.
[56] S.D. Khan, L. Alarabi, S. Basalamah, Toward smart lockdown: A novel approach for COVID-19 hotspots prediction using a deep hybrid, Neural Netw. 9 (4) (2020) 99.
[57] M.F. Aslan, M.F. Unlersen, K. Sabanci, A. Durdu, CNN-Based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection, Appl. Soft Comput. 98 (2021) 106912, 2021/01/01.
[58] V.N.M. Aradhya, M. Mahmud, D.S. Guru, B. Agarwal, M.S. Kaiser, One-shot cluster-based approach for the detection of COVID–19 from chest X–ray images, Cogn. Comput. (2021) 2021/03/02.
[59] P.K. Sethy, S.K. Behera, P.K. Ratha, P. Biswas, Detection of coronavirus disease (COVID-19) based on deep features and support vector machine, 2020.
[60] P.G.J. Moutounet-Cartan, Deep convolutional neural networks to diagnose covid-19 and other pneumonia diseases from posteroanterior chest x-rays, 2020.
[61] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Mach. Learn. 46 (1–3) (2002) 389–422.
[62] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, in: Advances in Neural Information Processing Systems, 2015, pp. 91–99.
[63] K. Han, J. Guo, C. Zhang, M. Zhu, Attribute-aware attention model for fine-grained representation learning, in: Proceedings of the 26th ACM international conference on Multimedia, 2018, pp. 2040–2048.
[64] X. Song, A. Elazab, Y. Zhang, Classification of Mild Cognitive Impairment Based on a Combined High-Order Network and Graph Convolutional Network, Vol. 8, (2020) pp. 42816–42827.
[65] S. Parisot, et al., Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease, Vol. 48, 2018, pp. 117–130.
[66] S.K. Pal, S. Mitra, Multilayer perceptron, fuzzy sets, classifiaction, 1992.
[67] A. Liaw, M. Wiener, Classification and regression by randomforest, R News 2 (3) (2002) 18–22.
[68] J. Ye, J.-H. Chow, J. Chen, Z. Zheng, Stochastic gradient boosted distributed decision trees, in: Proceedings of the 18th ACM conference on Information and knowledge management, 2009, pp. 2061–2064.