BEYOND COVID-19 DIAGNOSIS: PROGNOSIS WITH HIERARCHICAL GRAPH REPRESENTATION LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Coronavirus disease 2019 (COVID-19), the pandemic that is spreading fast globally, has caused over 32 million confirmed cases. Apart from Reverse Transcription Polymerase Chain Reaction (RT-PCR), the chest Computed Tomography (CT) and X-ray are viewed as standard and efficient tools for diagnosis and disease progression monitoring. We propose a diagnosis and prognosis model based on graph convolutional networks. The chest CT scan of a patient, typically involving hundreds of sectional images in sequential order, is formulated as a densely connected weighted graph. A novel distance aware pooling is proposed to abstract the node information hierarchically, which is robust and efficient for densely connected graphs. The new method, combining graph convolutional networks and distance aware pooling, can integrate the information from all slices in the chest CT scans for optimal decision making. It leads to the state-of-the-art accuracy in the COVID-19 diagnosis and prognosis. Our method achieves 94.6% accuracy for diagnosis and has a 2.2% improvement compared to the baseline model on the same dataset. In addition, we can localize the most informative slices with disease lesions for COVID-19 within the sequence of chest CT images. The proposed model can produce visual explanations for the diagnosis and prognosis, making the decision more transparent and explainable, while RT-PCR only leads to the test result with no prognosis information. We believe that the prognosis analysis can help hospitals designate medical resources more efficiently and better support clinicians to decide the proper clinical treatment.

1 INTRODUCTION

Coronavirus disease 2019 (COVID-19), which was first identified in December 2019 in China, has resulted in an ongoing pandemic in the world. Accorting to (ref, update) , until September 24th, it has caused more than 32 million confirmed cases in 188 countries and territories with more than 979,000 casualties. To control the sources of infection and cut off the channels of transmission, rapid testing and detection are of vital importance. The Reverse Transcription Polymerase Chain Reaction (RT-PCR) is a widely-used screening technology and viewed as the standard method for suspected cases (Organization et al., 2020). However, this method highly relies on the required lab facilities and the diagnostic kits. In addition, as the sensitivity of RT-PCR is not high enough for early diagnosis(Ai et al., 2020)(Fang et al., 2020). To mitigate the limitation of RT-PCR, radiographic imaging tests such as Computed Tomography (CT) and X-rays have been widely used as effective complimentary methods. In the diagnosis of COVID-19. CT is able to illustrate characteristic lesions in the patients' lungs (Huang et al., 2020), (Chung et al., 2020). Additionally, CT has also been proved to be useful in the COVID-19 disease progression and the therapeutic efficacy evaluation(Rodriguez-Morales et al., 2020), (Liechti et al., 2020).
According to radiog (Mao et al., 2018).

This work provides three primary contributions:

-
-
-

2 RELATED WORK

2.1 AI-ASSISTED AND CT-BASED COVID-19 DIAGNOSIS

Because of the limitations of RT-PCR in the diagnosis of COVID-19, radiographic imaging tests are now widely applied as effective supplements. Meanwhile, due to the high error rate of X-ray procedure, CT images are used for more precise detection(Alizadehsani et al., 2020). However, since each CT scan includes many image slices, it requires much time and labour of the radiologists (Shoeibi et al., 2020). Furthermore, since the radiological appearances of COVID-19 are similar to other types of pneumonia, radiologists need to go through much training before they can achieve a high diagnostic accuracy (Shi et al., 2020).

Due to above reasons, AI-assisted and CT-based COVID-19 diagnostic systems are preferably demanded. Recently, researchers have developed several such systems which show great performance. For example, (Chen et al., 2020) use Unet++ (Zhou et al., 2018) to segment infectious areas. (Butt et al., 2020) develop a deep learning model to detect lesions from CT images, and then use 3DResNet to classify the images into COVID-19, influenza-A viral pneumonia or healthy groups.(Song et al., 2020) use the whole lung for diagnosis instead of only extracting the lesions. They develop a Details Relation Extraction neural network (DRE-Net) for predictons in each CT image, and then aggregate the results to give the patient-level diagnosis. (Wang et al., 2020) propose an AI system that can both diagnose COVID-19 patients and do the prognostic analysis. They use FPN network(Lin et al., 2017) for lung area segementation, DenseNet(Huang et al., 2017) for classification, and a Cox proportional hazard model(Wang et al., 2019) to predict the hospital-stay time of COVID-19 patients.

2.2 GRAPH NEURAL NETWORKS

Different types of graph neural networks have been proposed, which can be divided in spectral or non-spectral domain. In spectral domain, the Fourier transformation and graph Laplacian help to define the convolution filters (Bruna et al., 2013). By utilizing Chebyshev polynomials, Kipf & Welling (2016) simplify the filter of graph convolution, rendering a lay-wise propagation method. However, the generalization of spectral methods may not be ideal due to the variety of graphs (Bronstein et al., 2017). Non-spectral methods, on the other hand, focus on the local topology of nodes, directly working on graphs instead of Fourier domain. Methods proposed by Hamilton et al. (2017), Monti et al. (2017) and Veličković et al. (2017) aggregate nodes based on adjacent nodes when the next layer is created. This aggregation process, as mentioned by Gilmer et al. (2017), can be regarded as a message passing process. Therefore, node information iteratively concentrates from neighboring nodes.

2.3 POOLING METHODS

Pooling methods allow GNN to hierarchically aggregate nodes, obtaining and assembling local information of graphs. The major purpose of hierarchical pooling method is to use a locally based model to aggregate nodes in each layer, so that a higher level graph representation will be created (Lee et al., 2019). Self-attention based pooling method, as proposed by Mao et al. (2018), is implemented with temporal video frame data, locally obtaining weighted and fused feature sequences and shadowing average pooling method, which homogeneously considers feature information. Spectral pooling methods, such as the one proposed by Ma et al. (2019), focus on the application of eigen-decomposition to capture the graph information. However, since spectral pooling methods are computationally demanding, they may not be accessible to large graphs. Non-spectral pooling methods are scalable to large graphs and pay attention to the local structure of graphs. Adaptive Structure Aware Pooling (ASAP) (Ranjan et al., 2020) and DiffPool (Ying et al., 2018) resort to clustering techniques, aggregating nodes into different clusters and then choosing top clusters based on cluster ranking scores (Gao & Ji, 2019). However, ASAP method partially ignores the edge weight information when using hops to aggregate nodes, resulting in unstable convergence.

3 METHODOLOGY

We proposed a graph convolutional network-based diagnosis and prognosis method that utilizes chest CT scans. The chest CT scans of a patient have a sequential and hierarchical data structure. Five CT findings with various rates, including ground glass opacification (GGO), interlobular septal thickening, bronchiectasis, pleural thickening, and subpleural involvement are commonly associated with COVID-19. Thus, the relationship between slices possesses more information than the order of the slices. The adjacent ones contains same abnormality could be considered as one lesion. The slices containing the same type of lesion may not be continuous since the lesions are distributed in various lung parts. For instance, as illustrated in figure, an example should be shown here.

We proposed a distance aware pooling method to downsampling and learn a graph-level representations from the input node features. The node features refers to the original CT scans in this paper. We model the sequential slices of CT scans hierarchically by a graph convolutional network integrated with the proposed pooling method. The model gradually extracts information from the slice level to the patient level by graph convolution and pooling. Eventually, a higher level of representation is learned and further used for diagnosis and prognosis. The overall model structure is illustrated in the figure below, which included GCN, Pooling modules and a classifier. The graph convolution-based method can integrate the information from all slices in the chest CT scans for optimal decision making. The model could coarsely localize the most informative slices within the chest CT scan. Thus, the proposed model can produce visual explanations for the diagnosis and prognosis, making the decision more transparent and explainable. We argue that this method could assist radiologists during diagnosis and prognosis in the clinical setting.

3.1 PROBLEM STATEMENT

Let $\mathcal{G}$ ($\mathbb{V}$, $\mathbb{E}$) be a patient's CT scan graph, with $|\mathbb{V}| = N$ nodes and $|\mathbb{E}|$ edges, where $|\cdot|$ represents the cardinality of a set. For each $v_i \in \mathbb{V}$, $\boldsymbol{x}_i$ is the corresponding d-dimensional vector. Let $\boldsymbol{X} \in \mathbb{R}^{N\times d}$ be the node feature matrix, and $\boldsymbol{A}^{adj} \in \mathbb{R}^{N\times N}$ be the weighted adjacency matrix. The entry in $\boldsymbol{A}^{adj}$ is defined based on cosine similarity, which is

$$\boldsymbol{A}^{adj}_{i,j} = \frac{< \boldsymbol{x}_i, \boldsymbol{x}_j >}{\|\boldsymbol{x}_i\| \cdot \|\boldsymbol{x}_j\|}. \tag{1}$$

Define distance matrix $\boldsymbol{A}^{dis}$ as

$$\boldsymbol{A}^{dis} = \mathbb{1} - \boldsymbol{A}^{adj}, \tag{2}$$

where $\mathbb{1}$ is the matrix with all 1's. The definition of $\boldsymbol{A}^{dis}$ means that the more similar two nodes are, the shorter distance they have. In the construction of $\boldsymbol{A}^{adj}$, the diagonal entries are automatically 1's since $i = j$ is allowed. Therefore, in the following calculation involving $\boldsymbol{A}^{adj}$, there is no need to put weight in the diagonal entries again.

The graph $\mathcal{G}$ has a label y, classes indicating health, Chlamydia Pneumonia, and COVID19. Thus, generally speaking, the diagnosis and prognosis of COVID-19 is a task of graph classification in this paper. Given a training dataset $\mathbb{T} = \{(\mathcal{G}_1, y_1), \ldots, (\mathcal{G}_N, y_N)\}$, the goal is to learn a mapping $f : \mathcal{G} \to \boldsymbol{y}$, which classifies a graph $\mathcal{G}$ into a class of $\boldsymbol{y}$. The model are composed of two modules, f_1 including node convolution and feature pooling, and f_2 involving a multilayer perceptron classifier which determines the class of each graph. At each layer of f_1, node embeddings and cluster membership are learnt iteratively. The first module can be written as $f_1 : \mathcal{G} \to \mathcal{G}^p$, where $\mathcal{G}^p$ is the pooled graph with fewer nodes and hierarchical feature representation. The second module is $f_2 : \mathcal{G}^p \to \boldsymbol{y}$, which utilized the graph-level representation learnt to diagnose and prognosis patients. The two modules are integrated in an end-to-end fashion. That is to say that the cluster assignment is learnt merely based on the graph classification objective.

3.2 NODE CONVOLUTION AND FEATURE POOLING

3.2.1 NODE CONVOLUTION

Node convolution applies the graph convolutional network to obtain high-level node feature representation in the feature matrix $\boldsymbol{X}$. Although there are many other methods to construct convolutional

network, the method recommended by Kipf & Welling (2016) is effective enough, which is

$$\boldsymbol{X}^{(l+1)} = \sigma(\sqrt{\boldsymbol{D}^{(l)}}\boldsymbol{A}^{adj(l)}\sqrt{\boldsymbol{D}^{(l)}}\boldsymbol{X}^{(l)}\boldsymbol{W}^{(l)}) \tag{3}$$

where $\boldsymbol{D}^{(l)}$ is the diagonal degree matrix of $\boldsymbol{A}^{adj(l)} - \boldsymbol{I}$, and $\boldsymbol{W}^{(l)} \in \mathbb{R}^{d\times k}$ is learnable weight matrix at the l-th layer. Due to the application of feature pooling method, the topology of the graph will be changed at each layer. Therefore, the dimension of matrices involved will be changed correspondingly.

3.2.2 POOLING METHOD

In this part, we proposed an innovative pooling method, which includes graph based clustering and feature pooling. We outline the pooling module and illustrate how it is integrated into an end-to-end graph convolutional neural network structure. Empirically, it is shown to be more robust for densely connected graph. The overall structure of the pooling method is illustrated in figure 1.

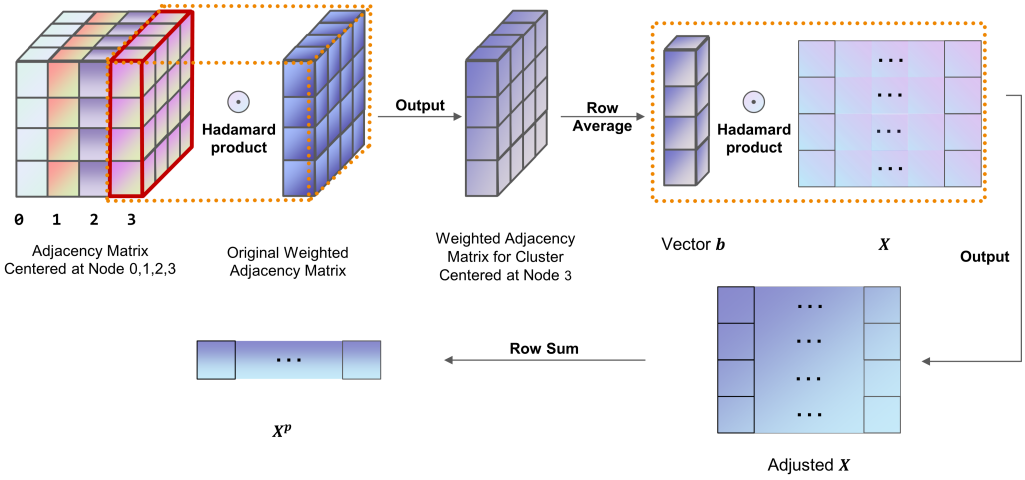

Figure 1: The structure of proposed distance aware pooling.

Distance aware Receptive Field The concept of receptive field RF used in convolutional neural network (CNN) was extended to graph neural network (GNN) (Ranjan et al., 2020). In Ranjan's work, they defined RF^{node} as the number of hops required to cover the neighborhood of a given node, such that given a chosen node, a cluster can be obtained based on a fixed receptive field h. However, this design may not be applicable to densely connected graphs. Because one node may be connected to most of the nodes in the graph. Even given a small value of h, for instance $h = 1$ which is the default in Ranjan's work, clusters formed within hop $h = 1$ may include most of the nodes of the graph. Hence, the cluster step may be bootless. Besides, the value of hops can only be integers, which loses the edge weight information in the clustering process. Therefore, we defined a distance aware receptive field for densely connected graphs RF^d denoted by h^d, which can be treated as a radius centered at a given node. The value of h^d is not restricted to integers but can be a positive real number.

Define $\mathcal{N}(v_i)$ as the local neighborhood of the node v_i, and $\mathcal{N}_{h^d}(v_i)$ as the RF^d neighborhood of the node v_i with a radius h^d. $\forall v_j \in \mathcal{N}(v_i)$, v_i and v_j are connected by the edge (v_i, v_j). However, $\forall v_p \in \mathcal{N}_{h^d}(v_i)$, there may exist another node v_q, such that v_p and v_i are connected by edges (v_i, v_q) and (v_q, v_p), and

$$A^{dis}_{i,q} + A^{dis}_{q,p} < h^d \tag{4}$$

Node clustering Inspired by the clustering and ranking ideas mentioned in (Ranjan et al., 2020) and (Gao & Ji, 2019), a local node clustering and score ranking method is proposed in the following. In the method, each node will be considered as a center of a cluster for a given h^d, then score all the clusters and choose the top k proportion of them to represent next layer's nodes with pooled feature values, where k is a hyper-parameter which needs to be determined.

Clustering ranking Given a node v_i and a radius h^d, $\mathcal{N}_{h^d}(v_i)$ is the corresponding RF^d neighborhood. Let $\mathbb{I}_1$ be the index set of the nodes in $\mathcal{N}_{h^d}(v_i)$. The cluster score is defined as

$$\alpha_i = \frac{\sum_{m,n\in \mathbb{I}_1} \boldsymbol{A}^{dis}_{m,n}}{|\mathcal{N}_{h^d}(v_i)|}, m \neq n. \tag{5}$$

It could be easily noticed that if α_i is a small value, nodes in $\mathcal{N}_{h^d}(v_i)$ are close to each other. Then, the top k proportion of clusters are to form next layer's nodes

Selecting cluster center $\forall v_j \in \mathcal{N}_{h^d}(v_i)$, define $\mathbb{V}_j = \mathcal{N}[v_j] \cap \mathcal{N}_{h^d}(v_i)$ as the set of nodes connected to v_j in $\mathcal{N}_{h^d}(v_i)$, where $\mathcal{N}[v_j]$ is the closed neighborhood of the node v_j. Let $\mathbb{I}_2$ be the index set of the nodes in $\mathbb{V}_j$. The node score is defined as

$$b_j = \frac{\sum_{c\in \mathbb{I}_2} \boldsymbol{A}^{dis}_{j,c}}{|\mathbb{V}_j|}. \tag{6}$$

The node with the smallest b_j value is chosen as the center of the cluster, and will be used to represent a new node in the next layer node representation.

Center node feature pooling Based on the node scores, rank the nodes in $\mathbb{V}_j$ and assign a weight w_r for $\boldsymbol{x}_r$, where $r \in \mathbb{I}_2$. Let $\hat{\boldsymbol{b}} = \mathbf{1} - \boldsymbol{b}$. The weight vector $\boldsymbol{w}$ is defined as

$$w_r = \frac{\hat{b}_r}{\sum\limits_{i\in\mathbb{I}_2} \hat{b}_i + \epsilon}, \tag{7}$$

where $\mathbf{1}$ is a vector with all 1's, ϵ is an extremely small positive value in order to avoid 0's in the denominator and $\boldsymbol{b}$ is the vector containing all the weights of nodes in $\mathbb{V}_j$. Then, the value of the pooled center feature is defined as

$$\boldsymbol{x}^p = \boldsymbol{X}_r \boldsymbol{w} \tag{8}$$

Then, $\boldsymbol{x}^p$ will be used as the center feature representing $\mathcal{N}_{h^d}(v_i)$.

Next layer node connectivity Following the idea in (Ying et al., 2018), the connectivity of the nodes in the next layer is preserved in the following method. According to above ranking and pooling methods, a pooled graph $\mathcal{G}^p$ with node set $\mathbb{V}^p$ has been obtained. The next step is to decide the the pooled adjacency matrix $\boldsymbol{A}^{adj,p}$. Define $\boldsymbol{S}$ such that the columns of $\boldsymbol{S}$ are the top k clusters' weight vectors $\boldsymbol{w}$. Hence, the pooled adjacency matrix is defined as

$$\boldsymbol{A}^{adj,p} = \boldsymbol{S}^T \boldsymbol{A}^{adj} \boldsymbol{S} \tag{9}$$

Then, the pooled distance matrix $\boldsymbol{A}^{dis,p}$ in the next layer is readily defined as

$$\boldsymbol{A}^{dis,p} = \mathbb{1} - \boldsymbol{A}^{adj,p}. \tag{10}$$

3.3 Graph Classification Based on Pooled Graphs

After finishing above processes, we will obtain the hierarchical representation for each patient i, which is a $N' \times D'$ matrix. In the results of each patient, N' represents those that are chosen to indicate the classification results. Besides, those N' CT slices included in the most informative cluster could provide a assistance for radiologists, helping them to identity potential abnormality in the lung.

4 EXPERIMENTS AND RESULTS

We evaluate our proposed method by comparing it with the state-of-art graph classification method for diagnosis and prognosis. Meanwhile, we aim to appraise whether the proposed methods compute meaningful and interpretable clusters on the input chest CT scans.

4.1 DATASET

We utilize the computed tomography (CT) database from the 2019 Novel Coronavirus Resource (2019nCoVR) (Zhang et al., 2020). The datasets include the complete chest CT scans of 929 COVID-19 positive patients, 964 Chlamydia Pneumonia patients, and 849 healthy people. The CT scans from 25 percent of patients are randomly chosen into the test set. Those from 60 percent of patients are assigned to the training set. The remaining CT scans are used for internal validation. It should be noted that the dataset is split according to patients instead of the CT scans. This is to avoid information leakage.

Moreover, the dataset provides these patients' clinical prognosis, whether the patients developed into severe/critical illness status. Severe/critical illness status refers to the admission to the intensive care unit (ICU), mechanical ventilation, or death. We believe that the prognosis analysis could support the hospital to designate medical resources more efficiently, and clinicians to decide the proper clinical treatment. Besides, the dataset summarizes the slices with lesions for COVID-19 assertive and Chlamydia Pneumonia patients. This is used to evaluate whether the cluster membership of the proposed method is interpretable.

4.2 DATA PREPROCESSING

- Standardize the size of chest CT scans
  Considering that each patient may have different sizes of chest CT images, our method first resizes each CT image into the same size, which in our method is denoted as $a \times b$.
- Standardize the number of chest CT scans of patients
  Since each patient have different number of CT scans, the method [method explanation required] is applied to appropriately choose the number of CT scans used in the downstream analysis.

4.2.1 CHEST CT SCAN IMAGE FEATURE EXTRACTOR

CNN Feature Extraction Utilize InceptionV3, ResNet which is pretrained on ImageNet Dataset for feature extraction.

Wavelet decomposition extractor Considering each CT scan as a 2-dimensional signal, we treat a scan image as a function with two variables and each image could be reconstructed as a summation of wavelet functions multiplying their coefficients for a given resolution (Mallat, 1989). The formula required to reconstruct CT scans is

$$\begin{aligned} f(x,y) =& \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(j_0,m,n)\varphi_{j_0,m,n}(x,y) \\ &+ \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{\infty} \sum_m \sum_n W^i_\Psi(j,m,n)\Psi^i_{j,m,n}(x,y). \end{aligned} \tag{11}$$

In above formula, $f(x,y)$ is the energy function of an image, M and N are the dimension of an image, W_φ and W_Ψ are treated as the coefficients of scaling function φ and wavelet function Ψ respectively. The coefficients W_φ are used in down streaming analysis since it contains the major approximation information of an image. The summation goes to infinity, and we only consider a finite number of cases. We choose the Haar wavelet function with resolution 3. The flattened approximation matrix of the image signal, which is a vector with dimension 1024×1 in our experiment, is used as the information vector of a CT scan in the calculation of image similarities.

4.3 QUALITATIVE RESULT

As illustrated in figurexx, the patient is diagnosed as COVID-19 positive.

4.4 QUANTITATIVE RESULT

- performance comparison table
- confusion matrix of diagnosis and prognosis
- performance as using some small portions of data for training.

Method	Feature Extractor	Accuracy	Precision	Recall	
Ours					

Table 1: Performance evaluation of COVID-19 Diagnosis

Method	Feature Extractor	Accuracy	Precision	Recall	F1 Score
Ours					

Table 2: Performance evaluation of COVID-19 Prognosis.

Method	Feature Extractor	No. Parameter	Inference Time
Ours			

Table 3: Computation Cost of COVID-19 Diagnosis and Prognosis methods.

5 CONCLUSION

REFERENCES

Tao Ai, Zhenlu Yang, Hongyan Hou, Chenao Zhan, Chong Chen, Wenzhi Lv, Qian Tao, Ziyong Sun, and Liming Xia. Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases. *Radiology*, pp. 200642, 2020.

Roohallah Alizadehsani, Mohaddeseh Behjati, Zahra Roshanzamir, Sadiq Hussain, Niloofar Abedini, Fereshteh Hasanzadeh, Abbas Khosravi, Afshin Shoeibi, Mohamad Roshanzamir, Pardis Moradnejad, et al. Risk factors prediction, clinical outcomes, and mortality of covid-19 patients. *medRxiv*, 2020.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203*, 2013.

Charmaine Butt, Jagpal Gill, David Chun, and Benson A Babu. Deep learning system to screen coronavirus disease 2019 pneumonia. *Applied Intelligence*, pp. 1, 2020.

Jun Chen, Lianlian Wu, Jun Zhang, Liang Zhang, Dexin Gong, Yilin Zhao, Shan Hu, Yonggui Wang, Xiao Hu, Biqing Zheng, et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study. *MedRxiv*, 2020.

Michael Chung, Adam Bernheim, Xueyan Mei, Ning Zhang, Mingqian Huang, Xianjun Zeng, Jiufa Cui, Wenjian Xu, Yang Yang, Zahi A Fayad, et al. Ct imaging features of 2019 novel coronavirus (2019-ncov). *Radiology*, 295(1):202–207, 2020.

Yicheng Fang, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, and Wenbin Ji. Sensitivity of chest ct for covid-19: comparison to rt-pcr. *Radiology*, pp. 200432, 2020.

Hongyang Gao and Shuiwang Ji. Graph u-nets. *arXiv preprint arXiv:1905.05178*, 2019.

Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. *arXiv preprint arXiv:1704.01212*, 2017.

Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in neural information processing systems*, pp. 1024–1034, 2017.

Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The lancet*, 395(10223):497–506, 2020.

Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Junhyun Lee, Inyeop Lee, and Jaewoo Kang. Self-attention graph pooling. *arXiv preprint arXiv:1904.08082*, 2019.

Marc R Liechti, Urs J Muehlematter, Aurelia F Schneider, Daniel Eberli, Niels J Rupp, Andreas M Hötker, Olivio F Donati, and Anton S Becker. Manual prostate cancer segmentation in mri: interreader agreement and volumetric correlation with transperineal template core needle biopsy. *European Radiology*, pp. 1–10, 2020.

Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117–2125, 2017.

Yao Ma, Suhang Wang, Charu C Aggarwal, and Jiliang Tang. Graph convolutional networks with eigenpooling. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 723–731, 2019.

Feng Mao, Xiang Wu, Hui Xue, and Rong Zhang. Hierarchical video frame sequence representation with deep convolutional graph network. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 0–0, 2018.

Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5115–5124, 2017.

World Health Organization et al. Clinical management of severe acute respiratory infection when novel coronavirus (2019-ncov) infection is suspected: interim guidance, 28 january 2020. 2020.

Ekagra Ranjan, Soumya Sanyal, and Partha P Talukdar. Asap: Adaptive structure aware pooling for learning hierarchical graph representations. In *AAAI*, pp. 5470–5477, 2020.

Alfonso J Rodriguez-Morales, Jaime A Cardona-Ospina, Estefanía Gutiérrez-Ocampo, Rhuvi Villamizar-Peña, Yeimer Holguin-Rivera, Juan Pablo Escalera-Antezana, Lucia Elena Alvarado-Arnez, D Katterine Bonilla-Aldana, Carlos Franco-Paredes, Andrés F Henao-Martinez, et al. Clinical, laboratory and imaging features of covid-19: A systematic review and meta-analysis. *Travel medicine and infectious disease*, pp. 101623, 2020.

Feng Shi, Jun Wang, Jun Shi, Ziyan Wu, Qian Wang, Zhenyu Tang, Kelei He, Yinghuan Shi, and Dinggang Shen. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. *IEEE reviews in biomedical engineering*, 2020.

Afshin Shoeibi, Marjane Khodatars, Roohallah Alizadehsani, Navid Ghassemi, Mahboobeh Jafari, Parisa Moridian, Ali Khadem, Delaram Sadeghi, Sadiq Hussain, Assef Zare, et al. Automated detection and forecasting of covid-19 using deep learning techniques: A review. *arXiv preprint arXiv:2007.10785*, 2020.

Ying Song, Shuangjia Zheng, Liang Li, Xiang Zhang, Xiaodong Zhang, Ziwang Huang, Jianwen Chen, Huiying Zhao, Yusheng Jie, Ruixuan Wang, et al. Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images. *medRxiv*, 2020.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Shuo Wang, Zhenyu Liu, Yu Rong, Bin Zhou, Yan Bai, Wei Wei, Meiyun Wang, Yingkun Guo, and Jie Tian. Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer. *Radiotherapy and Oncology*, 132: 171–177, 2019.

Shuo Wang, Yunfei Zha, Weimin Li, Qingxia Wu, Xiaohu Li, Meng Niu, Meiyun Wang, Xiaoming Qiu, Hongjun Li, He Yu, et al. A fully automatic deep learning system for covid-19 diagnostic and prognostic analysis. *European Respiratory Journal*, 2020.

Zhitao Ying, Jiaxuan You, Christopher Morris, Xiang Ren, Will Hamilton, and Jure Leskovec. Hierarchical graph representation learning with differentiable pooling. In *Advances in neural information processing systems*, pp. 4800–4810, 2018.

Kang Zhang, Xiaohong Liu, Jun Shen, Zhihuan Li, Ye Sang, Xingwang Wu, Yunfei Zha, Wenhua Liang, Chengdi Wang, Ke Wang, et al. Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography. *Cell*, 2020.

Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11. Springer, 2018.