A Review of Automated Diagnosis of COVID-19 Based on Scanning Images

Delong Chen[1], Shunhui Ji[1], Fan Liu[1,2] *, Zewen Li[1], Xinyu Zhou[3]

[1]College of Computer and Information, Hohai University
[2]Key Laboratory of Ministry of Education for Coastal Disaster and Protection, Hohai University
[3]China Pharmaceutical University
Nanjing, China

*fanliu@hhu.edu.cn

ABSTRACT

The pandemic of COVID-19 has caused millions of infections, which has led to a great loss all over the world, socially and economically. Due to the false-negative rate and the time-consuming of the conventional Reverse Transcription Polymerase Chain Reaction (RT-PCR) tests, diagnosing based on X-ray images and Computed Tomography (CT) images has been widely adopted. Therefore, researchers of the computer vision area have developed many automatic diagnosing models based on machine learning or deep learning to assist the radiologists and improve the diagnosing accuracy. In this paper, we present a review of these recently emerging automatic diagnosing models. 69 models proposed from February 14, 2020, to July 21, 2020, are involved. We analyzed the models from the perspective of preprocessing, feature extraction, classification, and evaluation. Based on the limitation of existing models, we pointed out that domain adaption in transfer learning and interpretability promotion would be the possible future directions.

Keywords

Deep learning, Machine learning, Biomedical Image Analysis, COVID-19.

1. INTRODUCTION

It has been six months since the first case of COVID-19 was confirmed. The pandemic was namely a cataclysmic catastrophe to the human. Up to July, the virus has registered over 10,000,000 infections and caused more than 50,000 death toll all over the world. In the battle between human and the novel coronavirus, early diagnosing and early quarantine is of vital importance. However, testing based on Reverse Transcription Polymerase Chain Reaction (RT-PCR) is time-consuming and may cause certain false-negative reports. To solve this problem, diagnosing based on scanning images (CT or X-ray) has been proved to be practical and effective. In the virus-stricken area, radiologists therefore have a heavy burden on analyzing scanning images. Since artificial intelligence based diagnosing models can assist the radiologists to reduce the diagnosing time and improve the accuracy, researchers have started to pay more attention to the development of COVID-19 diagnosing models. As shown in Fig. 1, the COVID-19 diagnosing models soared accordingly with the increasing of confirmed COVID-19 cases.

Due to the rapid development in this area, there have already been 9 reviews ([63]-[68], [92-94]) existing on this topic, but they have various shortcomings. On March 24, 2020, the first review of COVID-19 diagnosing models [63] critically argued the overfitting problem of existing models, but only six different modes were covered. Two days later, Joseph et al. [64] also discussed the same six diagnosing models and overviewed several artificial intelligence applications in the COVID-19 pandemic. Thanh [65] briefly summarized datasets, methods, and results of 12 diagnosing models. Shi et al. [66] and Muhammad et al. [67] focused on reviewing COVID-19 diagnosing models based on scanning images. Anwaar et al. [68] surveyed several computer vision applications for COVID-19, including diagnosis, prevention and control, clinical management and treatment. Mao et al. [92] surveyed data-driven analytical models for epidemic prediction, clinical diagnosis, policy effectiveness and contact tracing, but they covered only 5 artificial intelligence aided analysis models. However, to our best knowledge, there are at least 69 models that have been proposed, and many of them have not been covered by any of the existing surveys.

Recently, Chen et al. [93] presented a comprehensive survey on applications of artificial intelligence in fighting against COVID-19, including disease detection and diagnosis, virology and pathogenesis, drug and vaccine development, and epidemic and transmission prediction. They divided the diagnosing models into two categories: CT-based models and X-ray-based models, and present representative architecture for each of them. The representative architecture of CT-based models involves image segmentation, feature extraction, and classification and prediction, while the representative architecture of X-ray-based models involves model pre-training, fine-tuning, and prediction. But some X-ray-based models also consist of image segmentation and feature extraction process, and some CT-based models also consist of transfer learning process. On July 22, Afshin et al. [94] reviewed deep learning based diagnosing models, but machine learning is also an important branch of the research, which is not involved in their paper.

Therefore, as shown in Fig.2, in this paper we define a universal pipeline for diagnosing models, for both machine learning based models and deep learning based models. Then we organized the paper according to different stages in the model, rather than different types of scanning images used by the model. The contributions of this paper are as follows:

➤ We systematically reviewed and analyzed 69 COVID-19 diagnosing models from the perspective of preprocessing, feature extraction, classification, and evaluation. These models are proposed from February 14 to July 21, 2020.

➤ Based on the discussion of the existing models' limitation, we pointed out that domain adaption in transfer learning and interpretability promotion are the possible future directions.
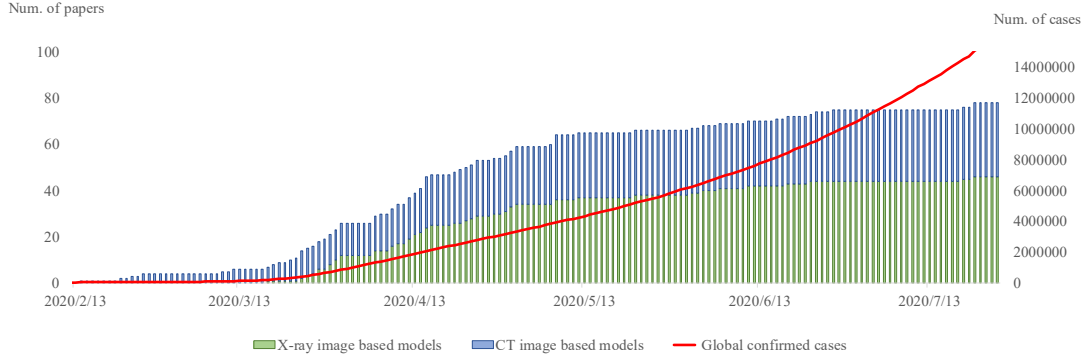
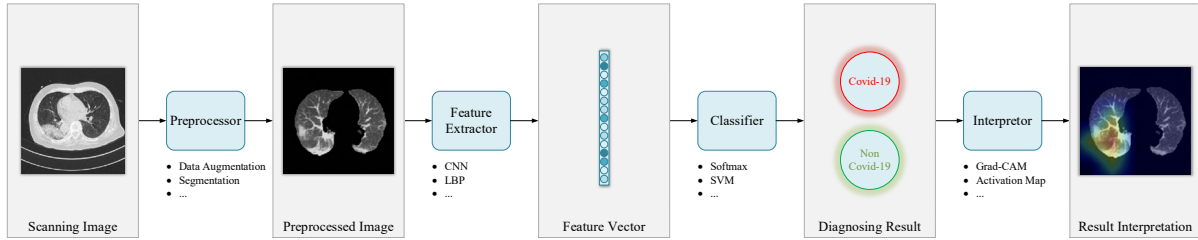Fig 1. Research of COVID-19 diagnosing models.

Fig 2. Typical pipeline of an automatic COVID-19 diagnosing model.

2. DIAGNOSIS METHODS OF COVID-19

Diagnosing COVID-19 based on scanning images is regarded as a classification task by most researchers. As shown in Fig.2, these classification models have a similar pipeline. First, the lung scanning images (CT or X-ray) are preprocessed, then the feature vectors are extracted by Convolution Neural Networks (CNN) or other feature extractors. The classifier predicts the infection. Finally, the model output heat map or bounding box to interpret its diagnosing result. In the following sections, we will review methods adopted in each stages of diagnosing models.

2.1 Preprocessing

In the existing literature, researchers have proposed three types of preprocessing methods: augmentation, equalization and segmentation. Data augmentation can enlarge the dataset and prevent overfitting, equalization improves the image quality, while lung segmentation can preserve the region of interest (ROI) only and avoid the undesired interfere from areas out of the lung.

To avoid overfitting, data augmentations are commonly adopted in the preprocessing stage. For simplicity, in Table I, we summarized basic transformation-based data augmentation methods used by COVID-19 diagnosing models. Among these basic methods, rotating, flipping, scaling, and cropping are the simplest and most common data augment methods, but their augment abilities are limited. In the comparative experiment of Mizuho et al. [#], the conventional data augmentation method improved the diagnosing performance by only 4%. Therefore, researchers proposed to apply other advanced data augmentation methods. Mehmet et al. [96] perform Zero-phase Component Analysis (ZCA whitening) to remove redundant information in input scanning images. Nour et al. [25] and Arvan et al. [40] used Generative Adversarial Network (GAN [46]) and Conditional Generative Adversarial Network (CGAN [47]) respectively to generate virtual samples for data augmentation. Generative models can significantly increase the dataset size, but the quality of the generated sample is difficult to guarantee. The purpose of data augmentation is to prevent overfitting by increasing the variation, but in these virtual sample methods, the discriminant lesion patterns might be lost or distorted if the model increases the variation too much.

Table I. Basic Transformation-based Data Augmentation Methods

Paper	Rotating or flipping	Scaling or cropping	Brightness adjusting	Contrast adjusting
[5]	√	√	-	-
[6]	√	√	√	√
[11]	√	√	-	-
[19]	√	√	-	-
[26]	√	√	-	-
[29]	√	√	√	-
[30]	-	-	-	√
[32]	√	-	-	-
[37]	√	√	-	√
[43]	√	-	-	-
[46]	√	√	-	-
[50]	√	√	√	-
[96]	√	-	-	-
[97]	√	√	-	-
[103]	-	√	-	-
[108]	√	√	√	√
Sum	14	12	4	4

To reduce the interference caused by different scanners and enhance the image contrast, Md et al. [31] and Oh et al. [34] performed histogram equalization on the images. But histogram equalization has potential harm of affecting image details or bringing unexpected noise. Md et al. [31] eliminated the noise by introducing Perona-Malikfilter (PMF) [77], while Lv et al. [52] and Manu et al. [104] solved the problem by proposing Contrast Limited Adaptive Histogram Equalization (CLAHE).

In both x-ray and CT images, areas out of the lung could interfere with the diagnosing model. Lung segmentation can reduce such undesired effects by preserving the region of interest (ROI) only. Lung segmentation can be carried out by radiologists [1], but it is time-consuming and inefficient. Morteza et al. [102] identify and remove the diaphragm region in X-ray images according to the brightness of the pixels, but such an algorithm is not robust to the interference caused by different scanners. U-Net [48] based methods were used in [5,6,8,28,38,41,51,52] for fast, automatic and accurate lung segmentation. Three models of them are X-ray scanning based models, but the other five of them are CT scanning based. For CT images, performing lung segmentation slice by slice will lose the contextual information between slices.

Therefore, some researchers [2,11] applied 3D versions of U-Net for lung segmentation, such as V-Net [78] and 3D U-Net++ [79]. Shan et al. further developed a fast VB-Net [72] by integrating the bottle-neck structure based on V-Net, they also adopted a human-in-the-loop training strategy to iteratively update the model and reduce the time-consuming annotation work. Based on this work, Shi et al. [10] and Sun et al. [53] adopted the VB-Net to segment the lung for location-specific feature extraction. Besides, other methods such as Dense-Net [80] in [13,34], OpenCV in [3], DeepLab in [9], and NABLA-N [32] were also adopted for lung segmentation.

2.2 Feature Extraction

Scanning images of COVID-19 has certain characteristic manifestations such as Ground Glass Opacity (GGO) and crazy-paving pattern distributed in a certain zone of lungs. Feature extraction is to detect those discriminative lesion patterns. Most COVID-19 diagnosing models adopted Convolutional Neural Network (CNN) for feature extraction, and most of them used existing network structures. We summarized some popular CNN structures that have been used by COVID-19 diagnosing models in Table II.

Table II. CNNs Used By COVID-19 Diagnosing Models

CNN Structure	Paper	Total
ResNet [81]	[2,3,5,7,8,9,11,12,16,17,18,19,20,24,25,28,31,33,34,35,42,43,44,45,48,49,50,52,97,99,100,107]	32
GoogLeNet [82]	[1,7,11,12,14,20,25,37,41,42,44,45,48,100,101,107]	16
DenseNet [80]	[3,12,19,31,36,41,43,44,45,50,51,52,54,99,100,107]	16
VGG [83]	[3,12,14,31,33,35,41,44,48,99,100,102,105,108]	14
MobileNet [84]	[12,14,27,36,41,44,50,99,107]	9
SqueezeNet [85]	[7,19,25,36,43,46]	6
AlexNet [86]	[7,15,19,23,25]	5
Capsule [87]	[26,40]	2

Some researchers also proposed automatic network structure designing methods. Wang et al. [22] used a generative synthesis approach to identify the optimal network architecture. Dalia et al. [30] applied Gravitational Search Algorithm (GSA) to determine the best network architecture hyperparameters. Sivaramakrishnan et al. [41] developed an iteratively pruning strategy to identify the optimal network structure.

Model ensemble can also promote the overall performance. Lawrence et al. [35] and Umut et al [29] performed model ensemble by voting and feature fusion. Md et al. [31] apply Softmax Class Posterior Averaging (SCPA) and Prediction Maximization (PM) for model ensemble, and Rodolfo et al. [39] combined seven traditional feature extraction models with Inception-v3 to obtain better results. Mahesh et al. [105] ensemble different CNNs by a stacked generalization approach [111] to further improve the model performance. These models assumed that different sub-models learn nonlinear discriminative features and semantic image representation from images of different levels. Therefore, the combined model will be more robust and accurate.

At the beginning of the pandemic, trying existing CNN is fast and convenient. However, these networks are designed for general image classification tasks such as ImageNet challenge [88]. Radiologists diagnose COVID-19 by finding distinguishing features such as local ground-glass appearance. Some researchers design local methods to extract more discriminative features. For

example, Umut et al. [29] and Oh et al. [34] used the local patches to train the CNN feature extractor. But lung infectious areas may vary significantly in size, the local methods with fixed patch size is unable to extract features of the target with the larger size. Hu et al. [38] proposed multi-scale learning to overcome such deficiency. The network aggregated features from different layers to make the final decision. Similarly, Ying et al. [3] integrated ResNet50 with the Feature Pyramid Network (FPN) [53], which is a pyramidal hierarchy network structure for multi-scale feature extraction. Besides, the lesion of COVID-19 in the lung is a 3D object, slice-wise contextual information in CT images would be lost by conventional 2D feature extractor. Therefore, Zheng et al. [6] proposed a CNN structure with 3D convolution units to detect COVID-19 to solve this defect.

In practice, radiologists also need to consider information such as epidemiology and clinical manifestations for diagnosis. Therefore, some methods also combine auxiliary external information with visual features to improve the model. Wang et al. [13] combine clinical features including age, sex, and comorbidity with CNN features. Similarly, since infected area usually lies near the edge of the lung, Xu et al. additionally provide the distance-from-edge information [2] of the local patch to the network. Shi et al. [10] and Sun et al. [53] calculate human-designed features including using volume, infection lesion number, histogram distribution, surface area, and radionics information

2.3 Classification

Classification is to present diagnosing prediction (such as COVID-19/non-COVID-19) according to the extracted feature. Most existing COVID-19 diagnosing models used CNN as the feature extractor, and most of them use softmax as the classifier. Some researchers proposed improvements based on the CNN + softmax scheme. For example, Wang et al. [1] combined softmax, decision tree, and Adaboost algorithms, Zhang et al. [16] simultaneously performed softmax loss based classification and contrastive loss-based anomaly detection to make the final decision. However, these deep models are black-box and usually need large-scale training sets. In literatures [7,21,29], researchers developed non-end-to-end models and taking Support Vector Machine (SVM) as the classifier. Comparative experiments of classification algorithms including SVM, logistic regression, k-Nearest Neighbors (k-NN), Multi-Layer Perception (MLP), decision tree, AdaBoost, random forest, LightGBM [89], and Bagging Classifier have been done in [10,39,44,53]. Among them, classifiers in [39] and [44] are for visual feature classification, while classifiers in [10] and [53] are for hand-craft clinical feature classification, in which Least Absolute Shrinkage and Selection Operator (LASSO) [90] and Deep Forest [91] algorithms were used for feature selection

A straightforward way of modeling the COVID-19 diagnosing task into a classification task is binary classifying the scanning images into COVID-19 class and normal class, and it is adopted by many models [4-6,9,12,16,20,21,23,27,29,32,40,43,44,46,49, 97,101,103,106,107]. But in practice, test images of other types of abnormal lung can be misclassified as COVID-19. As shown in Fig. 3 and Fig. 4, diagnosing COVID-19 is a fine-grained task. Lung diseases that belong to the same subclass share similar patterns in scanning images, and have a chance to be misclassified. Researchers overcome the problem of misclassification mainly through two approaches: multi-class classification and multi-step classification.

For multi-class classification, some researchers added other pneumonia categories in addition to the binary classification tasks, such as viral pneumonia, bacterial pneumonia, Community-Acquired Pneumonia (CAP), and non-COVID-19 pneumonia, as summarized in Table III. Other subclasses of viral pneumonia such as SARS, MERS are involved [15,39]. Some models also take account of other types of lung diseases, including ARDS [7,30], Tuberculosis [34], lung cancer [11,13], Pneumocystis, Streptococcus, Varicella [39], Fevers and upper respiratory tract symptoms [28].

However, methods of multi-class classification rely heavily on datasets. Meanwhile, the model cannot learn the hierarchical relationships between categories. Muti-step classification is to help the models to learn hierarchical relationships. For example, Eduardo et al. [33] and Yeh et al. [51] trained two binary classifiers, one for normal/pneumonia classification and one for further COVID-19/non-COVID-19 classification. Lv et al. [52] firstly classify the screening image into normal/bacterial pneumonia/viral pneumonia, then perform the COVID-19/non-COVID-19 classification. The above methods manually set up the hierarchical relationships for the models, while Rodolfo et al. [39] proposed to automatically learn a decision tree by the state-of-the-art Clus-HMC framework. They also comparatively tested multi-class classification and automatically multi-step classification (hierarchical classification), they reach a conclusion of the multi-step classification could be a feasible approach to improve COVID-19 recognition performance.
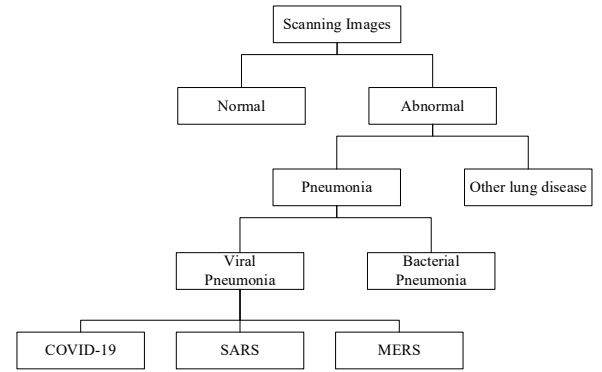

Fig 3. Hierarchical relationships of lung disease.
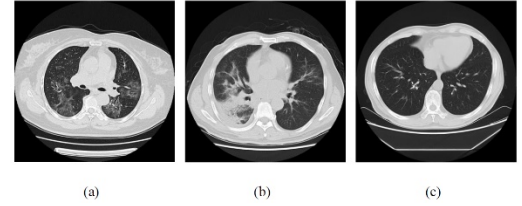
Fig 4. CT images of (a) COVID-19 (b) Influenza-A viral pneumonia (c) non-pneumonia.

Table III. Classes Involved in Multi-class Classification Models

	[2,18,104]	[3]	[17,22,24,26,31, 35,37,96,100,109]	[8,28]	[14,33,36,99, 102,108]	[1]	[10,95]	[25,98]
COVID19	√	√	√	√	√	√	√	√
Normal	√	√	√	√	√	-	-	-
Viral Pneumonia	√	-	√	-	-	√	-	-
Bacterial Pneumonia	-	√	√	-	-	-	-	-
Community-acquired pneumonia	-	-	-	√	-	-	√	
Non-COVID-19 pneumonia	-	-	-	-	√	-	-	√

Table IV. Dataset Size and Performance of COVID-19 Diagnosing Models

Date	Paper	Scanning type		Training set		Test set		Model performance	
		X-ray	CT	COVID-19	Total	COVID-19	Total	Accuracy	AUC
14-Feb	[1]	-	√	44	99	119	237	73.10%	78.00
21-Feb	[2]	-	√	219	618	30	90	86.70%	
23-Feb	[3]	-	√	53	165	26	82	94.00%	99.00
25-Feb	[4]	-	√	40	64	11	42	98.85%	
10-Mar	[5]	-	√	108	243	12	27		99.60
12-Mar	[6]	-	√	289	499	76	131	90.10%	95.90
18-Mar	[7]	√	-	80	160	27	54	95.30%	
19-Mar	[8]	-	√	400	3069	68	353		96.00
20-Mar	[9]	-	√	222	312	183	1072	94.98%	97.91
22-Mar	[10]	-	√	1326	2292	332	573	87.90%	94.20
23-Mar	[11]	-	√	723	1136	154	282		99.10
24-Mar	[12]	√	-	20	40	5	10	83.00%	90.00
24-Mar	[13]	-	√	1266	5372	102	226	85.00%	88.00
25-Mar	[14]	√	-	202	1284	22	143	97.82%	
26-Mar	[15]	√	-	74	137	32	59	95.12%	94.15
27-Mar	[16]	√	-	70	100	50	764	96.00%	95.18
27-Mar	[17]	√	-	54	4753	14	1188	89.82%	
28-Mar	[18]	-	√	100	513	50	256	98.80%	
29-Mar	[19]	√	-	130	2876	60	180	98.30%	99.80
29-Mar	[20]	√	-	40	80	10	20	98.00%	
30-Mar	[21]	√	-	25	40	unclear	unclear	97.48%	
30-Mar	[22]	√	-	76	16756	10	210	92.40%	
31-Mar	[24]	√	-	76	16756	10	210	96.23%	
31-Mar	[23]	√	√	144	263	144	263	98.00%	
2-Apr	[25]	√	-	unclear	624	unclear	1173	99.30%	
6-Apr	[26]	√	-	70	100	50	764	95.70%	97.00
6-Apr	[27]	√	-	409	3514	46	391	99.18%	
6-Apr	[28]	-	√	829	1865	109	199		99.40
7-Apr	[29]	-	√	2250	4500	750	1500	98.27%	
9-Apr	[30]	√	-	70	114	15	31	98.00%	
9-Apr	[31]	√	-	181	11896	78	5099	92.60%	
10-Apr	[32]	√	√	3875	5216	unclear	45	98.78%	
12-Apr	[33]	√	-	66	16546	10	210	92.80%	
12-Apr	[34]	√	-	126	354	36	99	88.90%	
13-Apr	[35]	√	-	122	410	33	242	94.40%	96.90
13-Apr	[36]	√	-	125	375	36	108		96.50
14-Apr	[37]	√	-	256	1125	28	126	89.50%	
14-Apr	[38]	-	√	120	360	30	90	89.20%	92.30
15-Apr	[39]	√	-	63	802	300	342	F1-score = 0.89	
15-Apr	[41]	√	-	286	14997	27	1703	99.01%	99.72
15-Apr	[40]	-	√	286	625	47	105		96.10
16-Apr	[42]	√	-	149	3783	31	11302	99.56%	
20-Apr	[43]	√	-	31	2031	40	3040		99.6
22-Apr	[44]	√	-	137	274	unclear	unclear	99%	
23-Apr	[45]	√	-	88	881	11	108	98%	99
24-Apr	[47]	√	-	180	718	45	175	95.30%	
24-Apr	[46]	-	√	251	1768	108	203	83%	
27-Apr	[48]	√	-	191	1791	48	448	97.01%	
29-Apr	[49]	√	-	258	2799	60	945		98.5
30-Apr	[50]	√	-	175	12092	20	1509	89.40%	
30-Apr	[51]	√	-	152	13594	31	231	96.80%	

Date	Ref							Accuracy	AUC
1-May	[52]	√	-	105	5486	10	405	83.12%	
7-May	[53]	-	√	1196	2018	299	504	91.79%	96.35
8-May	[54]	√	-	370	5029	92	1257	95.90%	
6-May	[95]	-	√	1047	1765	449	757	93.90%%	
8-May	[96]	√	-	370	5029	92	1257	95.90%	
12-May	[97]	√	-	147	4086	37	1202	95.50%	
22-May	[98]	√	-	101	131	152	197		90.06
1-Jun	[99]	√	-	195	998	20	125	83.60%	
3-Jun	[100]	√	-	189	989	47	247	F1-score = 0.89	
6-Jun	[101]	√	-	1000	50000	565	22594		94
11-Jun	[102]	√	-	415	8474	42	848	94.00%	
16-Jun	[103]	-	√	314	671	35	75	86.80%	3.1
18-Jun	[104]	√	-	429	1458	107	365	97.12%	
22-Jun	[105]	√	-	189	1935	56	555	92.74%	
23-Jun	[106]	-	√	657	3285	266	1330		
26-Jun	[107]	-	√	35	75	315	675	90.61%	96.05
19-Jul	[108]	√	-	398	5614	100	740	87.30%	
21-Jul	[109]	√	√	400	4685	50	1171	99.80%	
Average	-	-	-	354	3804	96	1059	93.36%	92.19

2.4 Evaluation

Researchers evaluated their proposed models with several metrics in experiments. The most used metrics are accuracy and the Area Under Curve (AUC). In Table IV, we summarized the evaluation of the above-mentioned COVID-19 diagnosing models. We also listed the submission date and the size of the training set and test set. For papers that not explicitly declared the size of training and test set, we calculate them with corresponding train/test set split ratio.

The average accuracy and AUC of diagnosing models based on X-ray scanning are 94.76% and 96.94, and the average accuracy and AUC of CT scanning based models are 90.13% and 94.76. Theoretically, 3D CT scanning contains more information than 2D X-ray scanning, and CT scanning can also avoid the occlusion of ribs compared with X-ray scanning. However, X-ray scanning based models achieve better performance. We consider the reason is the large size of X-ray training sets helps these models, while CT scanning is relatively more difficult to collect. The average training set size of X-ray based models is 4185, and the average size of CT based models is only 1417.

Although the performance of existing models is relatively high (average accuracy of 93.59% and average AUC of 95.75) but the size of the test set is worth noting, in some models, the test set has only a few COVID-19 samples. The average COVID-19/Total ratio of test sets is 0.274:1, which is highly imbalanced. (the ratio of training sets is 0.3:1, which is also imbalanced) In Table 4 we also color the table cells according to the number of samples in training and testing dataset (green and red are respectively corresponding to higher and lower than the average, and the saturation is correlated with the difference from the average). Some researchers reproduced the experiment with different dataset, but got significantly lower performance compared to the original reported performance [110]. The reason might be model overfitting and the lack of appropriate control patients and ground truth labels. Moreover, models in Table 4 are evaluated in different datasets, and most of them are private datasets. We think a proper benchmark test set is vital for further research of this area.

3. RESEARCH TREND ANALYSIS

3.1 Transfer Learning

A typical solution to the lack of massive scanning data is transfer learning. Among existing works, 34 models adopted the transfer learning scheme. They pre-trained the CNN on a larger image dataset (mostly on ImageNet), then fine-tune the model with X-ray or CT scanning images. But ImageNet contains images of general objects, which make the convolution filters to learn some patterns that will not appear in scanning images. Therefore, researchers proposed to transfer the model that pre-trained on lung cancer dataset [13] or conventional pneumonia dataset [51]. As another way to avoid overfitting, some researchers restricted or even skipped the fine-tuning of the CNN feature extractor. Wang et al. [1], Shervin et al [43], Narinder et al [45], and Sanhita et al. [47] only fine-tune certain part of the network, and Sethy et al. [7] and Ioannis et al. [27] skipped the fine-tuning process, only trained a classifier based on the features extracted by fixed CNNs. However, these methods can reduce the chance of overfitting to some extent, but they did not unleash the full potential of deep models.

Domain adaptation [56] is a branch of transfer learning, it's a learning technique to address the problem of lacking massive amounts of high-quality, large-scale labeled data. Fine-tuning only a certain part of the network can be regarded as domain adaptation [47], but there are also some specially designed deep domain adaptation models. Zhang et al. [49] used a domain discriminator to help the model better adapt to the target task. At present, there are few domain adaptation methods. we think applying deep domain adaptation to solve the problem of lacking massive training data of COVID-19 scanning images is an effective approach and valuable research direction.

3.2 Interpretability

In existing works, Class Activation Mapping (CAM) or Gradient-weighted Class Activation Mapping (Grad-CAM) [55] are adopted by researchers [5,8,9,16-18,22,28,31,34,36,38,41,45-47,49,51,52] to output heatmaps for explaining the final result and present an intuitive understanding of which area is the model

focusing on. At the same time, heatmaps can also provide radiologists with more useful information and further help them.

Compared to classification models, detection models can directly output a bounding box or a binary mask, they have an inherent advantage on interpretability. Detection based diagnosing is an emerging research direction. To the best of our knowledge, there are 5 detection based models [55-59] and 3 classification + detection models [60-62]. Radiologists search for lesions in the scanning images to make the diagnosis, so an object detection task can better simulate the human diagnosing process. Moreover, detection based methods can also avoid the information loss of local lesion patterns caused by the low dimension of the feature vector.

4. Conclusion

In this paper, we reviewed 69 automatic COVID-19 diagnosing models that emerged from February 14 to July 21, 2020. These models are based on machine learning or deep learning. They share a similar pipeline: preprocessing, feature extraction, classification, and evaluation. In the preprocessing stage of these models, transformation-based data augmentation and lung segmentation are performed. For feature extraction, most models adopted existing CNN structures, while others developed local methods to obtain more discriminative features. To enhance the performance of the classifiers, researchers proposed multi-class classification and multi-step classification. We also summarized the evaluation results of existing diagnosing models. Some of the models claim to perform well, but the size of test sets is not large enough.

Based on the limitations of existing models, we pointed out two possible future directions. Many existing models applied transfer learning to overcome the small dataset problem, but the adopted networks are pre-trained on general datasets such as ImageNet. To better utilize the information from both the source domain and target domain, a feasible solution is deep domain adaption. Besides, interpretability promotion is also important because it can further assist radiologists by providing more useful information. Detection based models have an inherent advantage on interpretability, such improvement towards better interpretability is also a valuable research direction.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Wang S., Kang B., Ma J., et al.: A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv.

[2] Xu X., Jiang X., Ma C., et al.: Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia. arXiv:2002.09334.

[3] Ying S., Zheng S., Li L., et al.: Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images. medRxiv.

[4] Chen J., Wu L., Zhang J., et al.: Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography a prospective study. medRxiv.

[5] Gozes O., Frid-Adar M., Greenspan H., et al.: Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic Initial Results for Automated Detection & Patient Monitoring using De. arXiv:2003.05037.

[6] Zheng C., Deng X., Fu Q., et al.: Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label. medRxiv.

[7] Sethy P. K., Behera S. K., et al.: Detection of Coronavirus Disease (COVID-19) Based on Deep Features.preprints 202003. 0300.v1.

[8] Li L., Qin L., Xu Z., et al.: Artificial Intelligence Distinguishes COVID-19 from Community-Acquired Pneumonia on Chest CT. Radiology, 2020: 200905.

[9] Jin C., Chen W., Cao Y., et al.: Development and Evaluation of an AI System for COVID-19 Diagnosis. medRxiv.

[10] Shi F., Xia L., Shan F., et al.: Large-Scale Screening of COVID-19 from Community-Acquired Pneumonia using Infection Size-Aware Classification. arXiv:2003.09860.

[11] Jin S., Wang B., Xu H., et al.: AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system in four weeks. medRxiv.

[12] Hemdan E. E. D., Shouman M. A., Karar M. E.: COVIDX-Net A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. arXiv:2003.11055.

[13] Wang S., Zha Y., Li W., et al.: A Fully Automatic Deep Learning System for COVID-19 Diagnostic and Prognostic Analysis. medRxiv

[14] Apostolopoulos I. D., Bessiana T.: COVID-19: Automatic detection from X-Ray images utilizing transfer learning with convolutional neural networks. arXiv: 2003.11617.

[15] Asmaa A., Mohammed M. A., Mohamed M. G.: Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. arXiv:2003.13815.

[16] Zhang J., Xie Y., Li Y., et al.: COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection. arXiv:2003.12338.

[17] Biraja G., Allan T.: Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection. arXiv:2003.10769.

[18] Fu M., Yi S., Zeng Y., et al.: Deep Learning-Based Recognizing COVID-19 and other Common Infectious Diseases of the Lung by Chest CT Scan Images. medRxiv.

[19] Muhammad E. H., Tawsifur R., Amith K. et al.: Can AI help in screening viral and COVID–19 pneumonia? arXiv:2003.13145.

[20] Ali N., Ceren K., Ziynet P.: Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. arXiv:2003.13145.

[21] Lamia N. M., Kadry A. E., Haytham H. E., et al.: Automatic X-ray COVID-19 Lung Image Classification System based on Multi-Level Thresholding and Support Vector Machine. medRxiv.

[22] Wang L., Alexander W.: COVID-Net A Tailored Deep Convolutional Neural Network Design for Detection of

COVID-19 Cases from Chest Radiography Images arXiv:2003.09871.

[23] Halgurd S. M., Aras T. A., Kayhan Z. G., et al.: Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms. arXiv:2004.00038.

[24] Muhammad F., Abdul H.: COVID-ResNet A Deep Learning Framework for Screening of COVID19 from Radiographs. arXiv:2003.14395.

[25] Nour E. M. K., Mohamed H. N. T., Aboul E. H., et al.: Detection of Coronavirus (COVID-19) Associated Pneumonia based on Generative Adversarial Networks and a Fine-Tuned Deep Transfer Learning Model using Chest X-ray Dataset. arXiv:2004.01184.

[26] Parnian A., Shahin H., Farnoosh N., et al.: COVID-CAPS A CAPSULE NETWORK-BASED FRAMEWORK FOR IDENTIFICATION OF COVID-19 CASES FROM X-RAY IMAGES. arXiv:2004.02696.

[27] Ioannis D. A., Aznaouridis I. S., Tzani A. M., et al.: Extracting possibly representative COVID-19 Biomarkers from X-Ray images with Deep Learning approach and image data related to Pulmonary. arXiv:2004.00338.

[28] Gozes O., Frid-Adar M., Nimrod S., et al.: Coronavirus Detection and Analysis on Chest CT with Deep Learning. arXiv:2004.02640.

[29] Umut O., Saban O., Mucahid B.: Coronavirus (COVID-19) Classification using Deep Features Fusion and Ranking Technique. arXiv:2004.03698.

[30] Dalia E., Aboul H., Hassan A. E.: GSA-DenseNet121-COVID-19 a Hybrid Deep Learning Architecture for the Diagnosis of COVID-19 Disease based on Gravitational Search Optimization Algorithm. arXiv:2004.05084.

[31] Md. R. K., Till D., Dietrich R. S., et al.: DeepCOVID-Explainer: Explainable COVID-19 Predictions Based on Chest X-ray Images. arXiv:2004.04582.

[32] Md Z. A., M M S. R., Mst S. N., et al.: COVID-MTNet COVID-19 Detection with Multi-Task Deep Learning Approaches. arXiv:2004.03747.

[33] Eduardo J. S. L., Pedro L. S., Rodrigo S., et al.: Towards an Efficient Deep Learning Model for COVID-19 Patterns Detection in X-ray Images. arXiv:2004.05717.

[34] Oh Y., Park S., Ye J. C.: Deep Learning COVID-19 Features on CXR using Limited Training Data Sets. arXiv:2004.05758.

[35] Lawrence O. H., Rahul P., Dmitry B., et al.: Finding COVID-19 from Chest X-rays using Deep Learning on a Small Dataset. arXiv:2004.02060.

[36] Li X., Li C., Zhu D.: COVID-MOBILEXPERT ON-DEVICE COVID-19 SCREENING USING SNAPSHOTS OF CHEST X-RAY. arXiv:2004.03042.

[37] Asif I. K., Junaid L. S., Mudasir B., et al.: CoroNet: A Deep Neural Network for Detection and Diagnosis of COVID-19 from Chest X-ray Images. arXiv:2004.04931.

[38] Hu S., Gao Y., Niu Z., et al.: Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification from CT Images. arXiv:2004.06689.

[39] Rodolfo P., Diego B., Lucas O. T., et al.: COVID-19 IDENTIFICATION IN CHEST X-RAY IMAGES ON FLAT AND HIERARCHICAL CLASSIFICATION SCENARIOS. arXiv:2004.05835.

[40] Arvan M., Pietro A. C., Samira Z., et al.: Radiologist-Level COVID-19 Detection Using CT Scans with Detail-Oriented Capsule Networks. arXiv:2004.07407.

[41] Sivaramakrishnan R., Jen S., Philip O. A., et al.: Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-rays. arXiv:2004.08379.

[42] Mohammad R., Abolfazl A.: A New Modified Deep Convolutional Neural Network for Detecting COVID-19 from X-ray Images. arXiv:2004.08052.

[43] Shervin M., Rahele K., Milan S., et al.: Deep-COVID: Predicting COVID-19 From Chest X-Ray Images Using Deep Transfer Learning. arXiv:2004.09363.

[44] Sara H. K., Peyman H. K., Michal J. W., et al.: Automatic Detection of Coronavirus Disease (COVID-19) in X-ray and CT Images: A Machine Learning-Based Approach. arXiv:2004.10641.

[45] Narinder S. P., Sonali A.: Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks. arXiv:2004.11676.

[46] Matteo P., Luigi C., Giuseppe P.: A Light CNN for detecting COVID-19 from CT scans of the chest. arXiv:2004.12837.

[47] Sanhita B., Sushmita M., Nilanjan S.: Deep Learning for Screening COVID-19 using Chest X-Ray Images. arXiv:2004.10507.

[48] Li T., Han Z., Wei B., et al.: Robust Screening of COVID-19 from Chest X-ray via Discriminative Cost-Sensitive Learning. arXiv:2004.12592.

[49] Zhang Y., Niu S., Qiu Z., et al: COVID-DA: Deep Domain Adaptation from Typical Pneumonia to COVID-19. arXiv:2005.01577.

[50] Brian D. G., Corey J., Can Z., et al.: Intra-model Variability in COVID-19 Classification Using Chest X-ray Images. arXiv:2005.02167.

[51] Yeh C., Cheng H., Wei A.: A Cascaded Learning Strategy for Robust COVID-19 Pneumonia Chest X-Ray Screening. arXiv:2004.12786.

[52] Lv D., Qi W., Li Y.: A cascade network for Detecting COVID-19 using chest x-rays. arXiv:2005.01468.

[53] Sun L., Mo Z, Yan F., et al.: Adaptive Feature Selection Guided Deep Forest for COVID-19 Classification with Chest CT. arXiv:2005.03264.

[54] Mehmet Y., Mete A., Aysen D., et al.: Convolutional Sparse Support Estimator Based Covid-19 Recognition from X-ray Images. arXiv:2005.04014.

[55] Zhou M., Chen Y., Yang D., et al.: Improved deep learning model for differentiating novel coronavirus pneumonia. medRxiv.

[56] Chen X., Yao L., Zhang Y.: Residual Attention U-Net for Automated Multi-Class Segmentation of COVID-19 Chest CT Images. arXiv:2004.05645.

[57] Zhou T., Canu S., Ruan S.: An automatic COVID-19 CT segmentation based on U-Net with attention mechanism. arXiv:2004.06673.

[58] Yu Q., Yun L., Jing X.: MiniSeg: An Extremely Minimum Network for Efficient COVID-19 Segmentation. arXiv:2004.09750.

[59] Fan D., Zhou T., Ji G., et al.: Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Scans. arXiv:2004.14133.

[60] Wu Y., Gao S., Mei J., et al.: JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation. arXiv:2004.07054.

[61] Xi O., Jiayu H., Liming X., et al.: Dual-Sampling Attention Network for Diagnosis of COVID-19 from Community-Acquired Pneumonia. arXiv:2005.02690.

[62] Tahereh J., Morteza H., Zohreh A., et al.: CovidCTNet: An Open-Source Deep Learning Approach to Identify Covid-19 Using CT Image. arXiv:2005.03059.

[63] Wynants L., Van Calster B., Bonten M. M. J., et al.: Systematic review and critical appraisal of prediction models for diagnosis and prognosis of COVID-19 infection. medRxiv.

[64] Bullock J., Pham K. H., Lam C. S. N., et al.: Mapping the landscape of artificial intelligence applications against COVID-19. arXiv:2003.11336.

[65] Nguyen T. T.: Artificial Intelligence in the Battle against Coronavirus (COVID-19): A Survey and Future Research Directions.

[66] Shi F., Wang J., Shi J., et al.: Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. IEEE Reviews in Biomedical Engineering, 2020.

[67] Ilyas M., Rehman H., Naitali A.: Detection of Covid-19 From Chest X-ray Images Using Artificial Intelligence: An Early Review. arXiv:2004.05436.

[68] Anwaar U., Asim K., Douglas G., et al.: COMPUTER VISION FOR COVID-19 CONTROL A SURVEY. arXiv:2004.09420.

[69] Goodfellow, I., Pouget-Abadie J., Mirza M., et al.: Generative adversarial nets in Proc. NIPS, 2014, pp. 2672-2680.

[70] Mirza M., Osindero S.: Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.

[71] Ronneberger O., Fischer P., Brox T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer International Publishing, 2015.

[72] Shan F., Gao Y., Wang J., et al. Lung Infection Quantification of COVID-19 in CT Images with Deep Learning. arXiv: 2003.04655.

[73] Lin T., Dollár P., Girshick R. et al. Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.

[74] Z.-H. Zhou and J. Feng, "Deep forest: Towards an alternative to deep neural networks," in Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 2017, pp. 3553-3559.

[75] Selvaraju R R., Das A., Vedantam R., et al. Grad-CAM: Why did you say that [J]. 2016.

[76] Wang M, Deng W. Deep visual domain adaptation: A survey. Neurocomputing, 2018, 312: 135-153.

[77] V Kamalaveni, R Anitha Rajalakshmi, and KA Narayanankutty. 2015. Image denoising using variations of Perona-Malik model with different edge stopping functions. Procedia Computer Science 58 (2015), 673-682.

[78] Field B , Znati T F , Mosse D . V-NET: a framework for a versatile network architecture to support real-time communication performance guarantees[C]// Infocom 95 Fourteenth Joint Conference of the IEEE Computer & Communications Societies Bringing Information to People IEEE. IEEE, 1995.

[79] iek, zgün, Abdulkadir A , Lienkamp S S , et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation[J]. 2016.

[80] Huang G , Liu Z , Laurens V D M , et al. Densely Connected Convolutional Networks[J]. 2016.

[81] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[J]. 2015.

[82] Szegedy C , Liu W , Jia Y , et al. Going Deeper with Convolutions[J]. 2014.

[83] Simonyan K , Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer ence, 2014.

[84] Howard A G , Zhu M , Chen B , et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[J]. 2017.

[85] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size[J]. arXiv preprint arXiv:1602.07360, 2016.

[86] Krizhevsky A , Sutskever I , Hinton G E . ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84-90.

[87] Sabour, S., Frosst, N., & Hinton, G.E. (2017). Dynamic Routing Between Capsules. ArXiv, abs/1710.09829.

[88] Russakovsky O , Deng J , Su H , et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3):211-252.

[89] Meng, Qi. (2018). LightGBM: A Highly Efficient Gradient Boosting Decision Tree.

[90] Tibshirani R . Regression shrinkage and selection via the lasso[J]. Journal of the Royal Statal Society, Series B, 1996, 58(1).

[91] Zhou Z H , Feng J . Deep Forest: Towards An Alternative to Deep Neural Networks[J]. 2017.

[92] Ying M., Susiyan J., Daniel N., et al.: Data-driven Analytical Models of COVID-2019 for Epidemic Prediction, Clinical Diagnosis, Policy Effectiveness and Contact Tracing: A Survey. arXiv:2006.13994.

[93] Chen J., Li K., Zhang Z., et al.: A Survey on Applications of Artificial Intelligence in Fighting Against COVID-19. arXiv:2007.02202.

[94] Afshin S.i, Marjane K., Roohallah A., et al.: Automated Detection and Forecasting of COVID-19 using Deep Learning Techniques: A Review. arXiv:2007.10785.

[95] Kang H., Xia L, Yan F., et al.: Diagnosis of Coronavirus Disease 2019 (COVID-19) with Structured Latent Multi-View Representation Learning. arXiv:2005.03227.

[96] Mehmet Y., Mete A., Aysen D., et al.: Convolutional Sparse Support Estimator Based Covid-19 Recognition from X-ray Images. arXiv:2005.04014.

[97] Sampa M., Seungwan J., Seiyon L., et al.: Multi-Channel Transfer Learning of Chest X-ray Images for Screening of COVID-19. arXiv:2005.05576.

[98] Zhou J., Jing B., Wang Z. et al.: SODA: Detecting Covid-19 in Chest X-rays with Semi-supervised Open Set Domain Adaptation. arXiv:2005.11003.

[99] Mizuho N., Shunjiro N., Hidetoshi M., et al.: Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods. arXiv:2006.00730.

[100] Soumick C., Fatima S., Chompunuch S., et al.: Exploration of Interpretability Techniques for Deep COVID-19 Classification using Chest X-ray Images. arXiv:2006.02570.

[101] Germán G., Aurelia B., José M., et al.: UMLS-ChestNet: A deep convolutional neural network for radiological findings, differential diagnoses and localizations of COVID-19 in chest x-rays. arXiv:2006.05274.

[102] Morteza H., Seyedehnafiseh M., Abolfazl Z., et al.: Improving performance of CNN to predict likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. arXiv:2006.12229.

[103] Chen X., Yao L, Zhou T., et al.: Momentum Contrastive Learning for Few-Shot COVID-19 Diagnosis from Chest CT Images. arXiv:2006.13276.

[104] Manu Siddhartha, Avik Santra: COVIDLite: A depth-wise separable deep neural network with white balance and CLAHE for detection of COVID-19. arXiv:2006.13873.

[105] Mahesh G., Sweta J.: Stacked Convolutional Neural Network for Diagnosis of COVID-19 Disease from X-ray Images. arXiv:2006.13817.

[106] Rohit L., Ashrika G., Viraj K., et al.: Automated Detection of COVID-19 from CT Scans Using Convolutional Neural Networks. arXiv:2006.13212.

[107] Abdolkarim S., Maryam S., Arash M. et al.: A Novel and Reliable Deep Learning Web-Based Tool to Detect COVID-19 Infection from Chest CT-Scan. arXiv:2006.14419.

[108] Zhong Y. et al.: Using Deep Convolutional Neural Networks to Diagnose COVID-19 From Chest X-Ray Images. arXiv:2007.09695.

[109] Md. K., Md. A., Md. T., et al.: CVR-Net: A deep convolutional neural network for coronavirus recognition from chest radiography images. arXiv:2007.11993.

[110] Imon B., Priyanshu S., Saptarshi P., et al.: Was there COVID-19 back in 2012? Challenge for AI in Diagnosis with Similar Indications. arXiv:2006.13262.

[111] D. H. Wolpert, Stacked generalization, Neural networks 5 (2) (1992) 241-259.