Identification and Localization COVID-19 Abnormalities on Chest Radiographs

van Tien Pham, Thanh Phuong Nguyen

Identification and localization COVID-19 abnormalities on chest radiographs *

Van Tien Pham[0000−0003−3890−7188]
Thanh Phuong Nguyen[0000−0002−5646−8505]

Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France
van-tien-pham@etud.univ-tln.fr; tpnguyen@univ-tln.fr

Abstract. Solutions to screen and diagnose positive patients for the SARS-CoV-2 promptly and efficiently are critical in the context of the COVID-19 pandemic's complex evolution. Recent researches have demonstrated the efficiency of deep learning and particularly convolutional neural networks (CNNs) in classifying and detecting lung disease-related lesions from radiographs. This paper presents a solution using ensemble learning techniques on advanced CNNs to classify as well as localize COVID-19-related abnormalities in radiographs. Two classifiers including EfficientNetV2 and NFNet are combined with three detectors, DETR, Yolov7 and EfficientDet. Along with gathering and training the model on a large number of datasets, image augmentation and cross validation are also addressed. Since then, this study has shown promising results and has received excellent marks in the Society for Imaging Informatics in Medicine's competition. The analysis in model selection for the trade-off between speed and accuracy is also given.

Keywords: chest X-ray · COVID-19 radiographs · classification and detection · deep neural network · medical imaging

1 Introduction

As the global coronavirus disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), spreads throughout the world, radiology is becoming increasingly important in giving clinical insights to help in illness diagnosis, treatment, and management. An increasing interest and use of chest X-ray (CXR) imaging has been demonstrated in a number of recent research [6], with some studies predicting a higher dependence on portable CXR and the high value of portable CXR for critically sick patients. CXR imaging systems are more widely available across the world than computed tomography (CT) scanners due to their lower cost and shorter decontamination periods. For identifying SARS-CoV-2, CXR imaging can be beneficial for patients with negative reverse transcription–polymerase chain reaction (RT-PCR) findings. As a result of the popularity of chest radiology imaging systems in current healthcare systems and the availability of portable equipment, radiography

* Thanh Phuong Nguyen is supported by project ANR ASTRID ROV-Chasseur.

examinations may be performed more quickly and with more availability, making this a good complement to RT-PCR testing [14]. Hence, CXR is an essential component of certain screening methods that have been presented. However, one of the most significant obstacles is the requirement for trained radiologists to interpret radiography imaging, as the visual indications might be subtle. While standards exist to assist radiologists in distinguishing COVID-19 from other forms of infection, radiologists' judgments vary. As a result, computer-aided diagnostic tools that can help radiologists analyze radiography pictures more quickly and reliably to discover COVID-19 patients are greatly sought.

Motivated by the urgent need of solutions to fight against the COVID-19 pandemic, this work tackles the problem of classification and detection COVID-19 abnormalities on chest radiographs via ensemble of deep CNNs. In particular, the radiographs will be categorized as 4 types: negative for pneumonia or typical, indeterminate, or atypical for COVID-19. Furthermore, the appearance of glass ground opacity and the extent to which they cover lung areas in the chest X-ray were also detected to help radiologists assess the COVID-19 pneumonia. Models from this research are evaluated in the competition [10] hosted by the Society for Imaging Informatics in Medicine, the Foundation for the Promotion of Health and Biomedical Research of Valencia Region and the Radiological Society of North America (SIIM-FISABIO-RSNA) on Kaggle. Our main contributions are:

- First, an in-depth exploratory data analysis is conducted on the SIIM-FISABIO-RSNA COVID-19 dataset [10] to reflect its properties and then suggest appropriate methods. Also, many open access benchmark datasets are collected and reviewed in order to pre-train models.
- Second, this work proposes a novel framework covering well-known object classifiers and detectors. Many methods, including transfer learning, data augmentation, and ensemble learning, are used to enhance the performance of models. This study also suggests using a classifier as a initial filter to eliminate false positive samples before passing to detectors. Separating the pulmonary approach proven to be an effective method concurrently.
- Third, several comparative evaluations are analyzed in terms of accuracy and computing cost. The recommendations for CXR-diagnosis practical applications are then highlighted.

2 Related works

Aim to interpret radiographic images better, many machine learning systems have been proposed [6], and the results show that the accuracy of radiographic imaging in detecting patients infected with COVID-19 is very promising while the cost of time is much reduced. Wang et al. [22] customized deep CNN models for identifying normal, pneumonia, and COVID-19 patients in the COVID-Net project. To provide systems aiding COVID-19 diagnosis, a variety of models [5] have been trained using X-ray and CT images. Shi et al. [15] retrieved a list of characteristics from CT images, including volume, histogram, and surface features. The open source COVID-19 Image Data Collection [4] is an example of

current initiatives to advocate for open access and solutions for radiography-driven COVID-19 case detection. Bilateral abnormalities, ground glass opacity, and interstitial abnormalities have all been identified in CXR radiograph of SARS-CoV-2 positive patients [7]. Similarly, in this study, we follow these concepts but also the rule of the SIIM-FISABIO-RSNA's competition with the 2 main tasks. For each test study in classification task, it's required to make a determination within 4 labels: *Negative for Pneumonia (NP), Typical Appearance (TA), Indeterminate Appearance (IA)* and *Atypical Appearance (AA)*. For each test image in detection task, a bounding box along with the *Opacity* class for findings and a corresponding confident score should be predicted. Otherwise, if there are no findings, the prediction should be *none 1 0 0 1 1* (*none* is the class ID for no finding, and this provides a one-pixel bounding box with a confidence of 1.0). We aim to accelerate the performance by updating the latest classifiers and detectors on huge datasets and further comparing with other researches in both performance and precision. Besides, many advanced techniques and tricks such as image augmentation, ensemble learning are implemented.

3 Proposed framework

This research proposes a framework for both classification and detection of COVID-19 abnormalities as illustrated in Fig. 1. For classification task, an input CXR which tagged as *study level* is inferred by the 4-classes classifiers. Meanwhile, the detection task is applied to CXRs marked as *image level*. At first, the input image is categorized into *Finding* or *No finding* by an EfficientNetV2 pre-trained as a 2-classes classifier to eliminate false positives. The *No finding* class indicates that there is no abnormality in this image, hence the ultimate prediction is *None*. On the contrary, the *Finding* sample is passed through a lung detector. Next, the dissected thorax is inferred by three detectors to find the location of the lesions which is labelled as *Opacity*. Finally, predicted bounding boxes and corresponding confident scores are then ensembled using weighted box fusion method to produce location of the lesion.

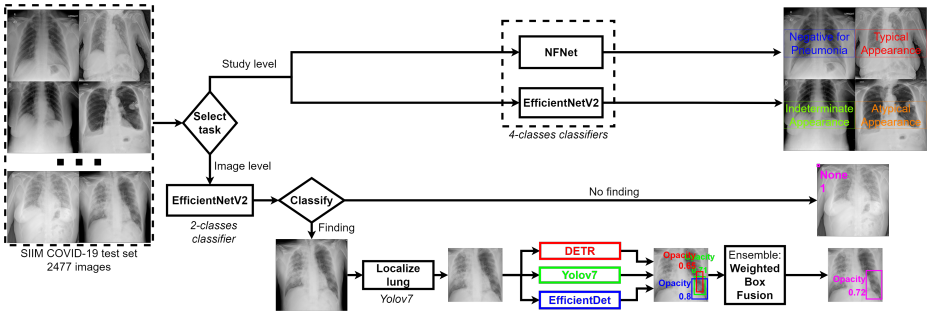

Fig. 1. Our framework for COVID-19 abnormalities classification and detection.

The training process of the models utilized in the framework is depicted in Fig. 2. The SIIM COVID-19 dataset is the primary dataset used in this study

for training and assessing models. To improve the generalization of the models, we also gather various external datasets. All data is preprocessed according to the method described in subsection 4.1, and image augmentation is performed as outlined in subsection 3.1. The models are trained through multiple iterations, with the best models selected through k-fold cross validation. These top models are then consolidated into a model zoo for use in the inference phase.

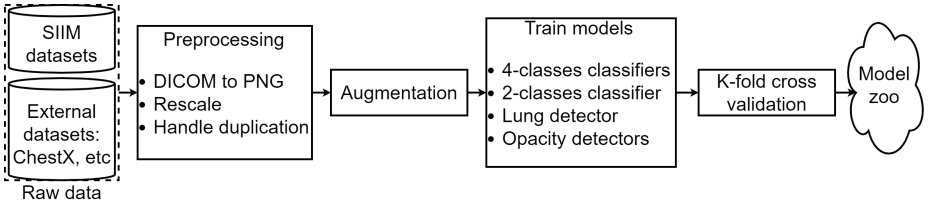

Fig. 2. Flow of training and validating models from raw data.

3.1 Data augmentation

Image augmentation is one of the key techniques used in deep learning and computer vision tasks to increase the quality and generality of trained models by creating fresh training samples from the existing data. The most prevalent augmentation strategies according to the expertise in medical image analysis are Blur, CLAHE, Cutout, ElasticTransform, GridDistortion, HorizontalFlip, IAASharpen, PiecewiseAffine, RandomResizedCrop, ShiftScaleRotate, and VerticalFlip. In this work, we adopted all of these method through the Albumentation library [2] while training models in the both tasks. Additionally, we apply Mosaic augmentation for detection task since this method not only modifies data on one image but also randomly combines data from four images into a new synthetic one as depicted in Fig. 3.

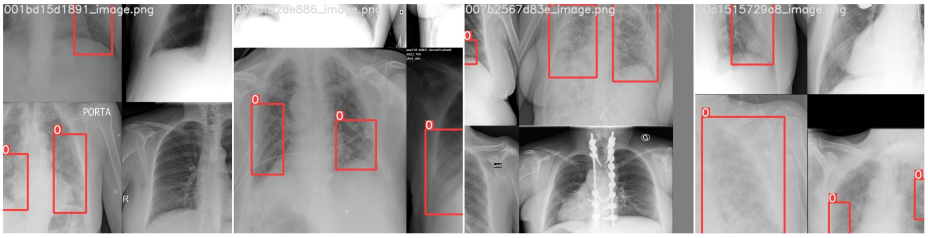

Fig. 3. Synthetic labels generated by Mosaic augmentation

3.2 COVID-19 abnormalities classifier

For classification task, two separate architectures, Normalizer-Free Networks (NFNet) [1] and EfficientNetV2 [17], are trained as 4-classes classifier to classify input image into 4 categories: *NP, TA, IA* and *AA*. NFNet is designed with the aim of overcoming instabilities for large learning rate or strong data augmentation with the proposal of the adaptive gradient clipping, which clips gradients based on the unit-wise ratio of gradient norms to parameter norms. A family

of Normalizer-Free ResNets from NFNet-F0 to NFNet-F6 are constructed and achieved higher validation accuracies on ImageNet while being faster to train. In this work, a NFNet-F5 is firstly pretrained on the CheXpert [8] and MIMIC-CXR [9] dataset and then fine-tuned on the main dataset at the image size of 512x512. EfficientNetV2 optimizes training time and parameter efficiency by combining training-aware neural architecture search and scaling. Crucial characteristics of the EfficientNetV2's backbone are: it heavily uses both MBConv and the newly introduced Fused-MBConv in the initial layers; it favors a lower kernel size but adds more layers; it eliminates the final stride-1 state employed in EfficientNet. During training, the regularization and image size are also adaptively changed as the concept of progressive learning to assist the network learn basic representations quickly and simply. Transfer learning technique are applied in this research on an EffcienNetV2-M which is pre-trained on the ChestX-ray14 dataset [23] with the aim of making it more general. Both of the two classifiers use a Binary Cross Entropy loss $\mathcal{L}_{BCE}$ blended loss of each class as in Eq. 1. The hyperparameters are selected empirically as follow: $\alpha = 0.2$, $\beta = 0.2$ and $\gamma = 0.3$ in order to reduce class imbalance phenomenon. Experiments show that this contrivance improves both cross-validation as well as final submission score.

$$L_{BCE} = \alpha L_{negative} + \beta L_{typical} + \gamma L_{indeterminate} + (1 - \alpha - \beta - \gamma) L_{atypical} \quad (1)$$

For detection task, instead of feeding chest radiographs directly to the object detectors, we first use a 2-classes classifier to filter out images that may contain abnormalities associated with the COVID-19 disease. As shown in Fig. 1, a EfficientNetV2-S is chosen as a classifier to separate inputs CXR into two categories *No finding* and *Finding*, corresponding to normal image and abnormal image respectively. This model is trained on these two classes extracted from the main dataset where we process *Negative for Pneumonia* labels as *No finding* and other labels as *Finding*. Initial weights are also pre-trained on the RICORD dataset [19] which contains 1000 CXRs rated by 3 different radiologists using the same 4-classes taxonomy as in the main dataset. This 2-classes classifier scheme brings two benefits: reduction of false positive rate and acceleration of speed during the inference phase. Because classifiers are frequently more accurate than detectors, this technique is extremely effective in filtering out early negative samples, especially for imbalanced datasets as observed in similar studies [13]. This rapid classifier also narrows down the CXRs that must be inferred on by the slower object detection models. For example, in the training dataset, there are 1736 *No finding* images out of 6334. This implies that more than a fourth of the entire data are eliminated, which leads to speed acceleration.

3.3 Lung detector

The purpose of lung detection is to separate between voxels that correspond to lung tissue and those that relate to surrounding anatomy. This localization serves to reduce background noise to assist the posterior detectors focus more on lung lesions, based on the assumption that nCoV virus-associated lesions are exclusively found in the lung area. A typical method for this task is the

simple center-crop at about 70% to 80% of the image. This approach proves to be effective and robust in cases where the lung is exactly in the center of the image. However, several factors influence this, including the capture angle, the radiologist's alignment, and the patient's standing position. When we examine the samples in this work's dataset, we see that the rib cage frequently deviates from the center of the picture. Therefore, a Yolov7 architecture is chosen instead of the center-crop method. 6334 CXRs in the training dataset are manually annotated to extract the lung bounding box. This data are then fed into training loops where a Yolov7 is employed to train on a single fold. Experiments show that this model boosts average precision while comparing with the simple center-crop.

3.4 COVID-19 lesion detector

The goal is to localize all findings, of the unique class *Opacity*, on a lung-cropped CXR with bounding boxes along with a confidence score. An example is shown as in Fig. 1 in which SOTA detectors (DETR, Yolov7 and EfficientDet) predict three different bounding boxes with score of 0.72, 0.65 and 0.78 respectively. DETR [3], combines a transformer encoder-decoder architecture with a set-based global loss that requires bi-partite matching to force unique predictions. Yolov7 [21] focuses on improving the accuracy without increasing the inference cost. EfficientDet [18] proposes a weighted bi-directional feature pyramid network for multiscale feature fusion and a compound scaling approach.

For implementation, DETR models are trained with ResNet-101 as backbone, an encoder-decoder transformer, a feed forward network and a learning rate of 0.001. Yolov7-E6 models are trained with E-ELAN as backbone and a learning rate of 0.005. EfficientDet-D7 model are also trained with SGD optimizer with a learning rate of 0.001 and EfficientNet as backbone. A Cosine Annealing scheduler and a Focal Loss function are applied. All backbone networks are pretrained on the RSNA dataset with *Opacity* class. Input image are scaled to 1024x1024 with a random augmentation described in 3.1 in training while during inference time, test time augmentation only consists of the horizontal and vertical flip. K-fold cross validation is employed to identify problems such as overfitting or selection bias, as well as to provide insight into how the models will generalize to an independent dataset, by splitting the training dataset into 5 folds based on patient's ID metadata extracted from DICOM data in the training stage.

3.5 Ensemble technique

The ensemble technique combines predictions of several models to achieve higher predictive performance than each of the constituent one could accomplish alone. These predictions consist of coordinates of the bounding box rectangle selected using non-maximum suppression based methods [11]. However, these approaches only produce boxes from a single model and can not adequately yield blended localization of predictions combined from various models. Therefore, we adopt the Weighted Boxes Fusion [16] in which coordinated of bounding rectangles as well as confidence scores from models are fused to construct the average boxes.

4 Experiments

4.1 Data collection and Evaluation protocol

Our proposed method is evaluated on different benchmark datasets. Tab. 1 summarizes data collection including 7 available CXR datasets from public sources. They are improved by augmentation techniques mentioned in subsection 3.1. In this work, a NFNet-F5 is firstly pretrained on the CheXpert dataset [8] and then fine-tuned on the main dataset at the image size of 512x512. Transfer learning technique are applied in this research on an EffcienNetV2-M which is pre-trained on the ChestX-ray14 dataset [23] in order to improve its generalization. Initial weights are also pre-trained on the RICORD dataset [19] which contains 1000 CXRs rated by 3 different radiologistis using the same 4-classes taxonomy as in the main dataset. The data is split in the ratio of 80:20 for the training and testing subsets. All images are in DICOM format and have been de-identified to safeguard patient privacy. A panel of qualified radiologists labels them based on the presence of opacity as well as the overall appearance.

Table 1. Description of collected CXR datasets.

Dataset	#Observations	#Image	#Patient	Annotation	Year
ChestX-ray14 [23]	15: No finding, 14 Pathologies	112,120	30,805	Classification	2017
MIMIC-CXR [9]	14: No finding, 12 Pathologies, Support devices	377,110	65,379	Classification	2019
CheXpert [8]	14: No finding, 12 Pathologies, Support devices	224,316	65,240	Classification	2019
VinDr-CXR [12]	28: No finding, 22 Lesions, 5 Diseases	18,000	18000	Detection	2020
BIMCV COVID-19+ [20]	20: Normal, COVID-19, Other findings	5381	1311	Classification	2020
RICORD [19]	4: Typical, Indeterminate, Atypical, Negative appearance for COVID-19	1,000	361	Classification	2021
SIIM [10]	4: Typical, Indeterminate, Atypical, Negative appearance for COVID-19 2: No finding, Opacity	7897	7268	Classification Detection	2021

All experiments are conducted on a RTX 1080. F1-score and mean Average Precision (mAP) at IoU 0.5:0.95 are employed as evaluation metrics.

4.2 Results and discussions

The experimental results of the classifiers are shown in Tab. 2. Covid-net [22] is selected as baseline model for comparing with the classifiers in this research. We modified the last fully connected layer and retrain the model to adapt with our datasets. Experiments show that the baseline model proved to be inefficient with 2 classes IA and AA, which are unavailable in it's original dataset. Our classifiers generally achieved better performance on the testing dataset.

Table 2. F1 score of the COVID-19 abnormalities classifiers on the SIIM test set

Classifier	F1 score			
	Negative for Pneumonia	Typical Appearance	Indeterminate Appearance	Atypical Appearance
Covid-net [22]	0.71	0.74	0.48	0.53
NFNet [1]	0.79	0.72	0.59	**0.56**
EfficientNetv2 [17]	**0.83**	**0.75**	**0.61**	0.54

The outcomes reveal a significant disparity between the four classes. When inferring the NP class, the EfficientNetv2 model earns the greatest F1 score of 0.83, while the NFNet model meets 0.79. The TA class achieved the second highest results, at 0.75 and 0.72 respectively for EfficientNetv2 and NFNet. The lowest result belongs to the case of the AA class, only 0.56 and 0.54 for

the two classifiers. The challenges come from the complexity of the "Atypical" class definition, which includes a range of abnormalities reported uncommonly for COVID-19 pneumonia such as pneumothorax or pleural effusion, pulmonary edema, lobar fusion, solitary pulmonary mass or nodule, diffuse small nodules and cavity. The model might not have enough samples of each abnormality to draw useful representations from given the variety of ways that unusual things might appear along with the dearth of data available for this class (10%). The features of the IA layer are also unclear, which leads to its low result, 0.59 and 0.61, respectively. There is a huge difference between the results of 2 classes NP and TA compared to 2 classes AA and IA. This is reasonable because symptoms on the NP and TA classes are often very specific and easily distinguishable, representing virus-positive and negative cases, respectively. In addition, the number of samples of these two classes also accounts for the majority of the test set.

Tab. 3 shows comparison between methods integrated in the proposed framework. If only the discrete models are counted, Yolov7 clearly outperforms the other two models with the best score of 0.563. A solid outcome was also achieved by the DETR architecture, which reached 0.542 accuracy points. When the EfficientDet model only managed to attain an accuracy of about 0.5, which is roughly 10% less than DETR and Yolov7, it became clear that it was not appropriate for this kind of situation. This can be accounted for by the lack of ample and diverse training data for a data-hungry architecture like EfficientNet.

Table 3. Evaluation of the COVID-19 lesion detector on the SIIM test set

Detector	Accuracy (mAP@ IoU 0.5:0.95)		Performance		
	Discrete model	Lung localized	Inference speed (FPS)	GPU memory requirement (MB)	Model size (MB)
DETR [3]	0.542	0.587	25	2683	232
Yolov7 [21]	**0.563**	0.591	**34**	3520	290
EfficientDet [18]	0.499	0.574	19	**1903**	**187**
Our Fusion	0.605	**0.612**	8	8106	709

The lung region identification and separation strategy proved to be very effective as it improved the accuracy for all models. For each of the three situations, the accuracy increased clearly, by 5%, 8%, and 15%, respectively. The most noteworthy of which is the application of lung separation by Yolov7 followed by the use of EfficientDet, which has produced a result of 0.499 to 0.591, an increase of 15%. The reason for this is most likely because the EfficientDet model, with the help of the lung partition model, has made it easier to focus on the core region of the lung. The case of overlapping two Yolov7 models proved to be ineffective when the results only increased by about 5%, this is because the architectural repetition in the two phases reduces the diversity of the results.

It should be noticed that the accuracy is significantly increased by employing the weighted box fusion technique, which merges predictions from three different models. It achieves an accuracy of up to 0.605, which is 7% higher than Yolov7 and a 21% increase over EfficientDet. The fusion solution reached an outstanding score of 0.612 in the case of previously recognized lung regions. However, as compared to the single findings, this outcome is only moderately improved by roughly 4%. This can be explained that the accuracy of the alternatives prior to compositing is already fairly high, so blending the results makes little sense.

Concerning performance evaluation, three factors are addressed. For inference speed, the Yolov7 architecture achieved the best results, 34 fps. DETR comes in second with 25 fps and EfficientDet comes in third with 19 fps. When all three models were performed in succession, the calculation speed dropped considerably to only 8 fps. In terms of GPU memory required when referencing the model, the best result belongs to EfficientDet with about 2 GB of GPU, while DETR (resp. Yolo7) consumes about 2.6GB (resp. 3.5 GB). If performing parallel models, the required memory is up to more than 8 GB. For memory space, EfficientDet needs only 187 MB, while DETR (resp. Yolov7) takes up 232 MB (resp. 290 MB).

The preceding results are used to make recommendations for the implementation of deep learning models in real-world scenarios. Diagnostic errors need to be kept to a minimum in both the general medical profession and COVID detection in particular. In order to merge these three models in a way that provides the best level of accuracy, it is important to employ the lung region separation technique and the weighted box fusion approach. This manner is much more expensive, it often results in a reduction of more than 4 times the computation speed when referring the model, increasing both the amount of GPU memory needed and the size of the storage. This approach is appropriate in cases where the accuracy is a priority, such as severe or highly suspected cases. On the other hand, we can use the Yolov7 straightforward model for a better trade-off between the accuracy and the consuming time. Though there may be a little, 5%, reduction in precision, there is a four-fold gain in calculating speed. The EfficientDet model can be used when the available hardware is constrained in terms of capacity because it consumes small resources to achieve satisfactory accuracy. Alternately, we can adopt the DETR model, which will provide a greater precision while only slightly increasing the amount of resources needed.

5 Conclusions

We have proposed a framework consisting of two classifiers and three detectors enhanced by ensemble learning approach to categorize and locate COVID-19 related anomalies. The results are thoroughly analyzed to provide recommendations in the application of deep learning models in practice. The limitations in data imbalance as well as the performance of the models will be improved in future researches. Our code is available at this address for research purpose.

References

1. Brock, A., De, S., Smith, S.L., Simonyan, K.: High-performance large-scale image recognition without normalization. arXiv preprint arXiv:2102.06171 (2021)
2. Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, M., Kalinin, A.A.: Albumentations: Fast and flexible image augmentations. Information **11**(2) (2020)
3. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S.: End-to-end object detection with transformers (2020)
4. Cohen, J.P., Morrison, P., Dao, L.: Covid-19 image data collection (2020). https://doi.org/10.48550/ARXIV.2003.11597, https://arxiv.org/abs/2003.11597

5. Costa, Y.M.G., Silva, S.A., Teixeira, L.O., Pereira, R.M., Bertolini, D., Britto, A.S., Oliveira, L.S., Cavalcanti, G.D.C.: Covid-19 detection on chest x-ray and ct scan: A review of the top-100 most cited papers. Sensors **22**(19) (2022)
6. Gomes, R., Kamrowski, C., Langlois, J., , C., Haley, M.: A comprehensive review of machine learning used to combat COVID-19. Diagnostics (Basel) p. 1853 (2022)
7. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al.: Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. The lancet **395**(10223), 497–506 (2020)
8. Irvin, J., Rajpurkar, P., Ilcus, S.: Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. CoRR **abs/1901.07031** (2019)
9. Johnson, A.E.W., Pollard, T.J., Berkowitz, S.J., Greenbaum, N.R., Lungren, M.P., Deng, C., Mark, R.G., Horng, S.: MIMIC-CXR: A large publicly available database of labeled chest radiographs. CoRR **abs/1901.07042** (2019)
10. Lakhani, P., Mongan, J., Singhal, C., Zhou, Q., Andriole, K.P., Auffermann, W.F., Prasanna, P., Pham, T., Peterson, M., Bergquist, P.J., et al.: The 2021 siim-fisabio-rsna machine learning covid-19 challenge: Annotation and standard exam classification of covid-19 chest radiographs. (Oct 2021)
11. Neubeck, A., Van Gool, L.: Efficient non-maximum suppression. In: 18th International Conference on Pattern Recognition (ICPR'06). vol. 3, pp. 850–855 (2006)
12. Nguyen, H.Q., Lam, K., Le, L.T., Pham, H.H., Tran, D.Q., Nguyen, D.B., Nguyen, A.T., Ho, P.H., Ngo, D.T., Nguyen, N.T., Nguyen, N.T., Dao, M., Vu, V.: Vindr-cxr: An open dataset of chest x-rays with radiologist's annotations (2020)
13. Pham, V.T., Tran, C.M., Zheng, S., Vu, T.M., Nath, S.: Chest x-ray abnormalities localization via ensemble of deep convolutional neural networks. In: ATC. pp. 125–130 (2021)
14. Rubin, G.D., Ryerson, C.J.: The role of chest imaging in patient management during the covid-19 pandemic: a multinational consensus statement from the fleischner society. Radiology **296**(1), 172–180 (2020)
15. Shi, F., Wang, J., Shi, J., Wu, Q., Tang, Z., He, K., Shi, Y., Shen, D.: Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for covid-19. IEEE Reviews in Biomedical Engineering **14**, 4–15 (2021)
16. Solovyev, R.A., Wang, W.: Weighted boxes fusion: ensembling boxes for object detection models. CoRR **abs/1910.13302** (2019)
17. Tan, M., Le, Q.V.: Efficientnetv2: Smaller models and faster training. CoRR **abs/2104.00298** (2021), https://arxiv.org/abs/2104.00298
18. Tan, M., Pang, R., Le, Q.V.: Efficientdet: Scalable and efficient object detection. CoRR **abs/1911.09070** (2019), http://arxiv.org/abs/1911.09070
19. Tsai, E., Simpson: The rsna international covid-19 open radiology database(ricord). Radiology **299**(1), E204–E213 (2021)
20. Vayá, M.d.l.I., Saborit, J.M., Montell, J.A., Pertusa, A., Bustos, A., Cazorla, M., Galant, J., Barber, X., Orozco-Beltrán, D., García-García, F., Caparrós, M., González, G., Salinas, J.M.: Bimcv covid-19+: a large annotated dataset of rx and ct images from covid-19 patients (2020)
21. Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M.: Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors (2022)
22. Wang, L., Lin, Z.Q., Wong, A.: Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports **10**(1), 1–12 (2020)
23. Wang, X., Peng, Y., Lu, L., Bagheri, M., Summers, R.M.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. CoRR **abs/1705.02315** (2017)