 doi:10.20944/preprints202203.0288.v1

TOWARDS CLINICAL PRACTICE: DESIGN AND IMPLEMENTATION OF CONVOLUTIONAL NEURAL NETWORK-BASED ASSISTIVE DIAGNOSIS SYSTEM FOR COVID-19 CASE DETECTION FROM CHEST X-RAY IMAGES

Daniel Kvak, MSc* • Marian Bendik • Anna Chromcova, MD

Carebot s.r.o., Rozdrojovice 26E, Czech Republic

Abstract

One of the critical tools for early detection and subsequent evaluation of the incidence of lung diseases is chest radiography. This study presents a real-world implementation of a convolutional neural network (CNN) based Carebot Covid app to detect COVID-19 from chest X-ray (CXR) images. Our proposed model takes the form of a simple and intuitive application. Used CNN can be deployed as a STOW-RS prediction endpoint for direct implementation into DICOM viewers. The results of this study show that the deep learning model based on DenseNet and ResNet architecture can detect SARS-CoV-2 from CXR images with precision of 0.981, recall of 0.962 and AP of 0.993.

Keywords: *computer-aided detection, convolutional neural network, COVID-19, deep learning, image classification.*

 doi:10.20944/preprints202203.0288.v1

1 Introduction

COVID-19 disease causes severe acute respiratory syndrome 2 (SARS-CoV-2), which was originally discovered in late 2019. On January 30, 2020, the outbreak was declared a Public Health Emergency of International Concern and subsequently, on March 11, 2020, WHO declared COVID-19 a pandemic. (Wu et al. 2020) Once infected, a patient with COVID-19 may develop various symptoms and signs of infection, which include fever, cough, and respiratory illness. In severe cases, infection can cause pneumonia, respiratory distress, multi-organ failure and death. (Cascella et al. 2022, Shereen et al. 2020, Rothan & Byrareddy 2020)

At a time when the speed and reliability of results, especially for COVID-19 positive patients, is important, the development of applications that would facilitate the work of untrained staff involved in the evaluation is also crucial. Chest radiography is an important tool for early detection and subsequent verification of lung diseases. (Qin et al. 2018, Speets et al. 2006) As such, radiographic testing can be performed more quickly and has greater accessibility due to the prevalence of chest radiological imaging systems in modern healthcare systems and the availability of portable units, making them a suitable adjunct to RT-PCR testing, particularly as CXR imaging is often performed as part of the standard procedure for patients with respiratory difficulties. (Martínez Chamorro et al. 2021, Kvak & Kvaková 2021) However, in the current pandemic situation of COVID-19, we encounter a lack of radiologists and trained staff to analyze the vast number of images taken. (Cavallo & Forman 2020)

Many researchers have published a series of preprints demonstrating computer-aided detection (CADe)-based approaches to detect COVID-19 from chest radiographs. (Apostolopoulos & Mpesiana 2020, Elgendi et al. 2020, Nahiduzzaman et al. 2021) These approaches have achieved promising results on a small dataset but are by no means production-ready solutions. The aim of our research is to present a real-world CNN-based approach that would eliminate the time and resources needed to develop new technologies and related algorithms. (Meenatchi Aparna & Shanmugavadivu 2018)

 doi:10.20944/preprints202203.0288.v1

Work	Dataset	Model	Accuracy	F1 Score
(Apostolopoulos et al. 2020)	455 COVID-19, 2109 Non-COVID Images	MobileNet V2	99.18%	-
(Mahmud et al. 2020)	305 COVID-19, 1888 Normal, 3085 Bacterial Pneumonia, 1798 Viral Pneumonia	Stacked MultiResolution CovXNet	97.4%	0.971
(Ozturk et al. 2020)	127 COVID-19, 500 Normal, 500 Pneumonia Images	DarkCovidNet	98.08%	0.965
(El Asnaoui & Chawki 2021)	231 COVID-19, 1583 Normal, 2780 Bacterial Pneumonia, 1493 Viral Pneumonia Images	Inception ResNetV2	92.18%	0.921
(Khan et al. 2020)	284 COVID-19, 310 Normal, 330 Bacterial Pneumonia, 327 Viral Pneumonia Images	CoroNet	99%	0.98

Table 1: Comparison with state-of-the-art methods.

2 Radiology perspective

The COVID-19 disease can be identified on a conventional chest radiograph based on several typical patterns. The two most common patterns are ground-glass opacities and lung consolidation. However, the ground-glass opacities observed on a chest CT can be difficult to detect on a chest radiograph, it is often accompanied by a reticular opacities region which is more easily appreciable on a standard CXR. (Cozzi et al. 2021) Lung consolidations in COVID-19 (and other viral pneumonia) can be often found multifocally, the distribution is usually bilateral and includes lower lobes. (Jacobi et al. 2020) One of the most specific features of COVID-19 on a CXR is a peripheral and posterior distribution of air-space opacities. In more serious stages of the disease when patients are typically hypoxic, diffuse air-space opacities covering the majority of lung parenchyma can be found and the CXR pattern can be similar to acute respiratory distress syndrome (ARDS). The rare or less common findings in COVID-19 are pleural effusions, pneumothorax, and lung cavitation. (Jacobi et al. 2020)

During the first four days of symptomatic COVID-19 disease, there can be a normal finding on the chest radiograph and chest CT scan. Later on, when there are present the signs of COVID-19 on CXR, the images can show a high similarity to those of several types of viral pneumonia and other inflammatory lung diseases. Therefore, it is difficult for medical doctors to distinguish COVID-19 infections from other viral pneumonia using only a chest X-ray. However, the conventional chest radiograph as well as the chest CT scan were found to be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

an important diagnostic tool in addition to PCR test for their higher sensitivity, availability, speed, and possible prediction of severity of the disease. (Tahir et al. 2021)

3 Proposed model architecture

In the recent past, deep learning has been very successful in a variety of visual tasks. Deep learning-based models have revolutionized CADe by accurately analyzing, identifying and classifying patterns in medical images. (Yamashita et al. 2018) In the past, deep learning has had success in mammography image classification. (Shen et al. 2019) The success of machine learning-based algorithms in CADe and the rapid growth of COVID-19 cases have necessitated the need for an automatic detection and diagnosis system based on artificial intelligence. Recently, many researchers have proposed the use of CNN-based CADe models to detect COVID-19 from CXR. (Apostolopoulos et al. 2020, Mahmud et al. 2020, Ozturk et al. 2020, El Asnaoui & Chawki 2021, Khan et al. 2020)

In recent years, the advent of a new generation of GPUs and the deployment of cloud environments such as Microsoft Azure and Amazon AWS have kick-started a second wave of innovation in CADe. (Kagadis et al. 2013) Unlike earlier approaches which utilized handcrafted features, the CNN-based models are trained to identify relevant biomarkers on a large dataset without explicit input from researchers.

3.1 Convolutional neural network

For image recognition and classification tasks, various CNN architectures have proven to be widely used. The basic idea is that neurons in the visual cortex process images into increasingly complex shapes. (Krizhevsky et al. 2012) The image is first segmented at the boundaries of edges using a light/dark interface, then fused into simple shapes, and finally fused into recognizable complex features in subsequent layers. (Krizhevsky et al. 2012) CNN tries to mimic this idea using multiple layers of artificial neurons. The standard architecture includes several convolutional layers that segment the image into small pieces that can be easily processed. (Albawi et al. 2017)

The outputs from these layers are aggregated into layers to reduce the size of the data and reduce noise. The sequential layers feed into a neural network, which then produces a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

Model	Layers	Parametres	Input matrix	Output activation
ResNet50V2 (He et al. 2016)	50	25.6M	224, 224, 3	softmax
DenseNet121 (Huang et al. 2017)	121	8M	224, 224, 3	softmax

Table 2: Architecture of the proposed models implemented in Carebot Covid app.

probabilistic heat map that describes the probability of whether the image contains the desired target. (Krizhevsky et al. 2012) The advantage of this system is that it can be trained to find any feature in an image without the designer having to describe specific features. (Yamashita et al. 2018) The convolutional layer is an important part of the deep learning neural network that extracts features from the input images. This avoids a major limitation of previous hand-crafted algorithms, which required hard-coding specific identifying features into the software.

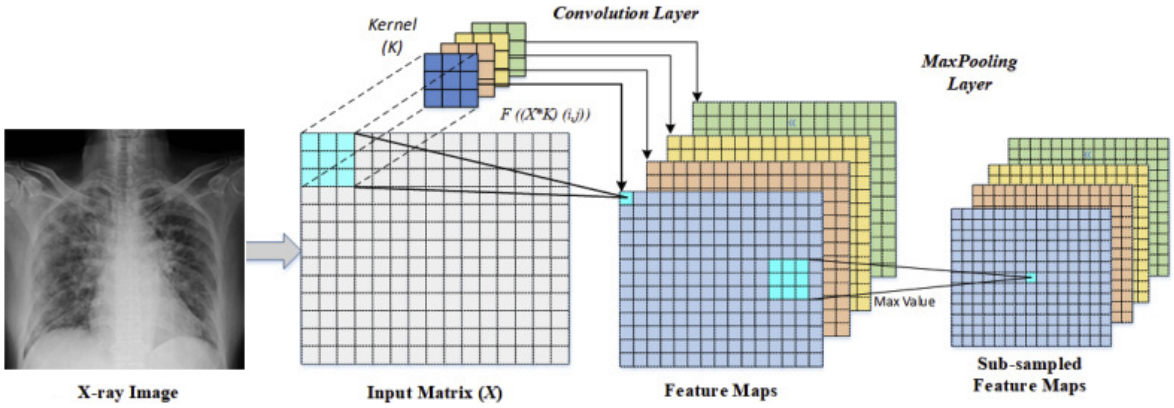

Figure 1. Schematic representation of convolution and max-pooling layer operations inside a CNN.

3.2 DICOMweb implementation

Current paradigms for the use of medical image resources call for vendor-neutral archives, accessible through standard interfaces and with support for multiple repositories. The communication processes, data format, storage, querying, retrieval, visualisation and printing of this medical image information are specified by the international standard Digital Imaging and Communications in Medicine (DICOM), which is the main standardisation effort in this field. (Lebre et al. 2020) Patient information is bundled in one or more standard files that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

contain metadata related to the patient, study, or report in addition to image pixels. A typical infrastructure consists of one or more archives, acquisition methods (data production units), distribution mechanisms, and visualization devices. (Lebre et al. 2020)

The DICOMweb STOW-RS (STore Over the Web) service is the basis of the IHE WIC integration profile. (Clunie et al. 2016) WIC seeks to simplify the sender's task by allowing not only the sending of DICOM objects, but also the transmission of non-DICOM format images or videos along with minimal DICOM metadata in XML or JSON encoding, and by shifting the burden of filling in the DICOM metadata describing the pixel data to the server. (Le Maitre et al. 2014)

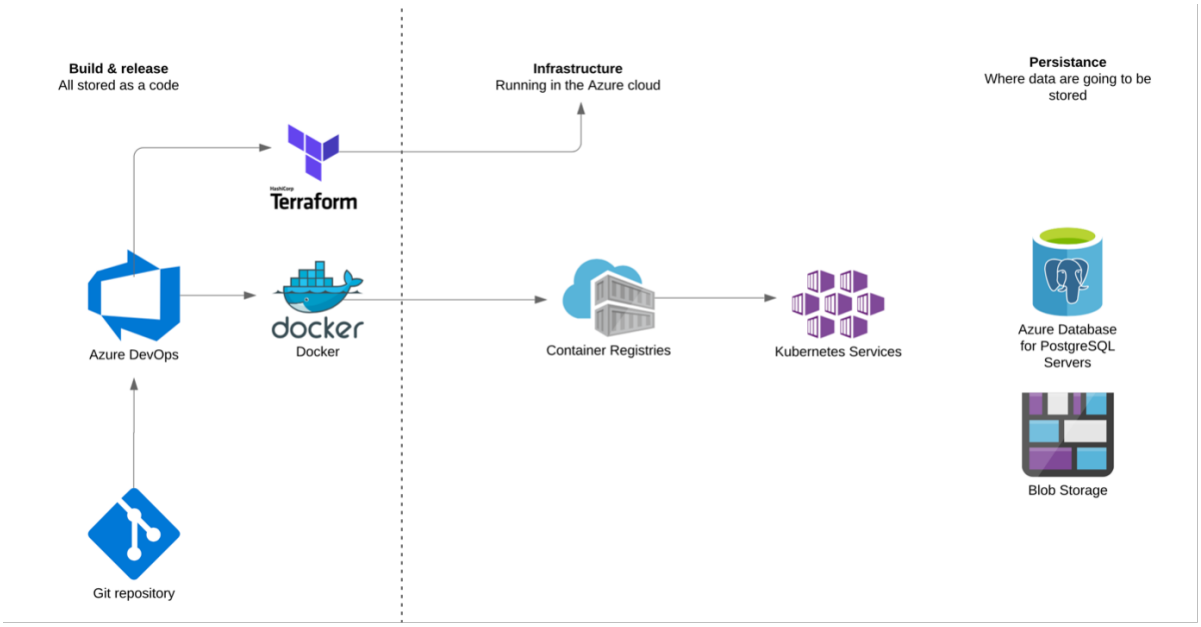

Figure 2. Production environment high-level architecture of Carebot Covid app.

STOW-RS has been defined to reduce the burden for implementing lightweight dispatch applications, for example on mobile devices. (Genereaux et al. 2018) This allows the sender to use an HTTP POST operation to send one or more DICOM objects, either encoded as regular DICOM PS3.10 binary files, or as separate XML or JavaScript Object Notation (JSON) metadata and separate image or video pixel data in various Internet media file formats. (Le Maitre et al. 2014)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

Communication in the opposite direction, i.e. from the prediction server back to the user, is handled by the DICOMweb WADO-RS (Web Access to Dicom Persistent Objects) middleware. The endpoint is integrated into the Picture Archiving and Communication Systems (PACS) web system and provides services for a specific PACS system. It can be deployed in a hospital to provide the WADO service to healthcare professionals and used in regional PACS to transfer medical images and messages. (Liu et al. 2015) The use of WADO-RS allows us to build a DICOM Structured Report object that contains the requested CADe application predictions.

3.3 Out-of-distribution detection

CNNs are often trained with a closed-world assumption, i.e., the distribution of the test data is assumed to be similar to the distribution of the training data. (Yamashita et al. 2018) However, when deployed in real-world, this assumption does not hold, leading to significant performance degradation. Although this performance degradation is acceptable for applications such as product recommendation, it is dangerous to use such systems in intolerant domains such as CADe because they can cause serious accidents. (Firmino et al. 2016) The proposed Carebot Covid app generalizes to out-of-distribution (OOD) examples whenever possible, flagging those that are beyond its capabilities and seeking human intervention. (Hendrycks & Gimpel 2016)

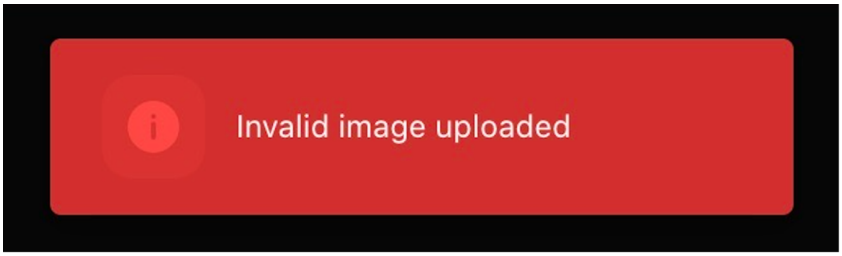

Figure 3. While similar solutions try to predict the diagnosis even from the inappropriate image, Carebot Covid app successfully detects that it is not a CXR image and discards it as invalid.

One of the difficult tasks is knowing when not to make a prediction. If you ask a radiologist to diagnose an image of something that is not his specialty, he should not provide a diagnosis. For CNN, its specialization is the classifier domain, which is defined by its training

 doi:10.20944/preprints202203.0288.v1

distribution. (Hermann et al. 2020) We cannot evaluate the accuracy of the model on these examples, so we choose not to process them, hence only images that are similar to the images in the training distribution can be processed. A CNN-based binary classifier method was selected for model design and calibration, which was trained on a subset of the original dataset and the Tiny ImageNet dataset. (Le & Yang 2015) An OOD model was built on this annotated dataset to predict whether new examples belong to the positive or negative class. (Hendrycks & Gimpel 2016)

4 Subjects and dataset

Selected images were obtained from publicly available datasets. The combination of different datasets increases the confidence in the developed identification models, and it also increased the size of the dataset, which is a problem in most of the related literature. (Apostolopoulos et al. 2020, Mahmud et al. 2020, Ozturk et al. 2020, El Asnaoui & Chawki 2021, Khan et al. 2020)

Thus, the full train and test dataset contains 21,905 unique CXRs, of which 3,987 represent CXR images of patients with a positive COVID-19 test, 7,650 CXR images show patients with findings unrelated to COVID-19 infection (pulmonary lesions, fibrosis, pneumothorax etc.), and 10,268 CXR images record patients with no or negligible pathological findings. CXR were taken in posteroanterior (PA) or anteroposterior (AP) projection in patients who were unable to stand. All images in AP projection were taken using portable X-ray machines with patients in supine or sitting position. (Zhang et al. 2021)

To create the dataset, we combined and modified several different publicly available datasets. Examples of CXR images from the dataset used are shown in Figure 1 and illustrate the diversity of patient cases (including age, sex, stage of infection, or imaging projection) in the dataset. Patients younger than 15 years were excluded from the dataset, as well as images of poor quality or incorrect projection.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

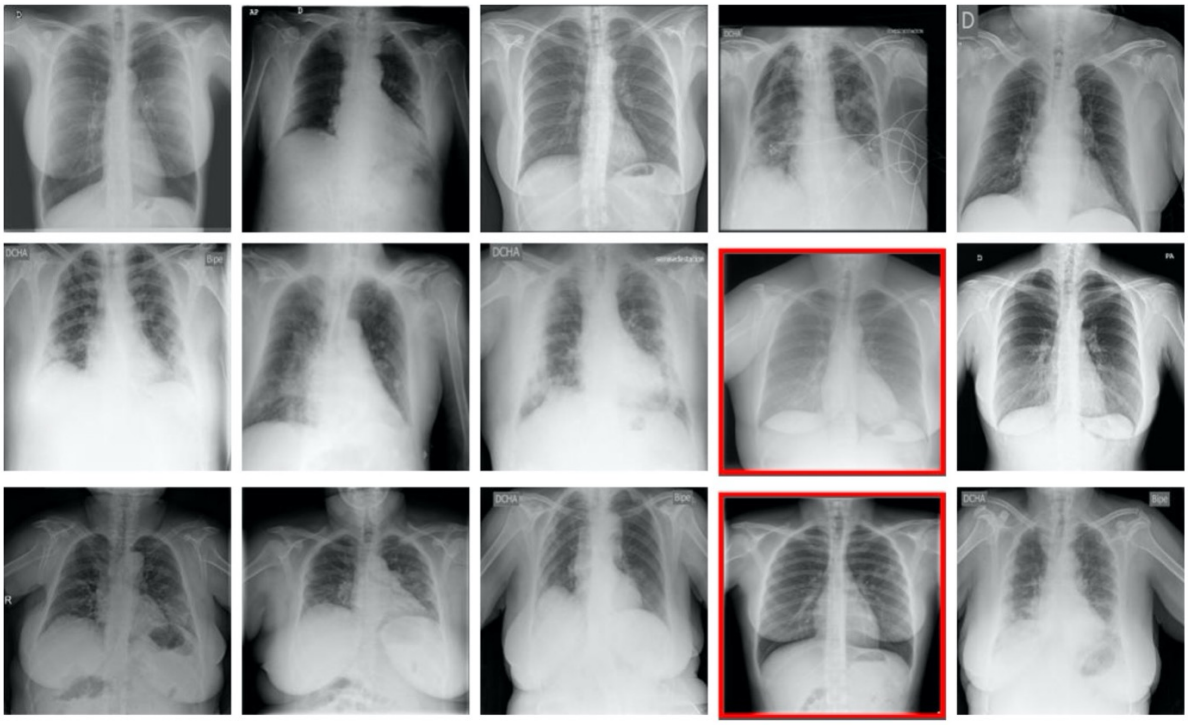

Figure 4. CXR test images during neural network training. Marked red are those for which CNN incorrectly predicted the class.

4.1 Data augmentation

Data augmentation increases the size of the input training data along with the regularization of the model, thus improving the generalization of the training model. (Mikołajczyk & Grochowski 2018) It also helps to create new train examples by randomly applying different transformations to the available dataset to reflect the noisiness of real-world data. (Shorten & Khoshgoftaar 2019, Elgendi et al. 2021) In our study, we used transformations involving vertical flipping of training images, random rotations, modifications in lighting conditions, zoom, saturation, and JPEG encoding noise. The training data was augmented using five randomly selected variations, the extension of test dataset was not investigated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

5 Classifier performance

An F1 Score becomes a critical evaluation tool to determine False Positive and False Negative rates yielded through a discriminating threshold in a similar situation with unbalanced dataset samples. (Sokolova et al. 2006) The classification performance of the Carebot Covid app model for multi-class problem was evaluated for each component and the average classification performance of the model was calculated. The following table includes the accuracy, precision, recall, and F1 Score, calculated based on the following equations below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

For specific experiments and given that there is a class imbalance problem, the most reliable metric is the model average accuracy metric, while given that this accuracy is high, the second most important metric is the recall metric for individual classes. (Japkowicz & Stephen 2002) This is due to the importance of correctly identifying true cases that are not COVID-19 (True Negatives). AP (Average Precision) summarizes a precision-recall curve as the weighted mean of precisions achieved at each threshold (Yilmaz & Aslam 2006), with the increase in recall from the previous threshold used as the weight:

$$\text{AP} = \sum_{n} (R_n - R_{n-1}) P_n \tag{5}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

Class	Image count	Precision	Recall	F1 Score	AP
Model Average	21,905	0.952	0.950	0.951	0.985
COVID-19	3,987	0.981	0.962	0.971	0.993
Non-COVID-19	7,650	0.952	0.922	0.937	0.980
No Finding	10,268	0.941	0.967	0.954	0.984

Table 3: Averaged test results of Carebot Covid app for multi-class classification after cross validation.

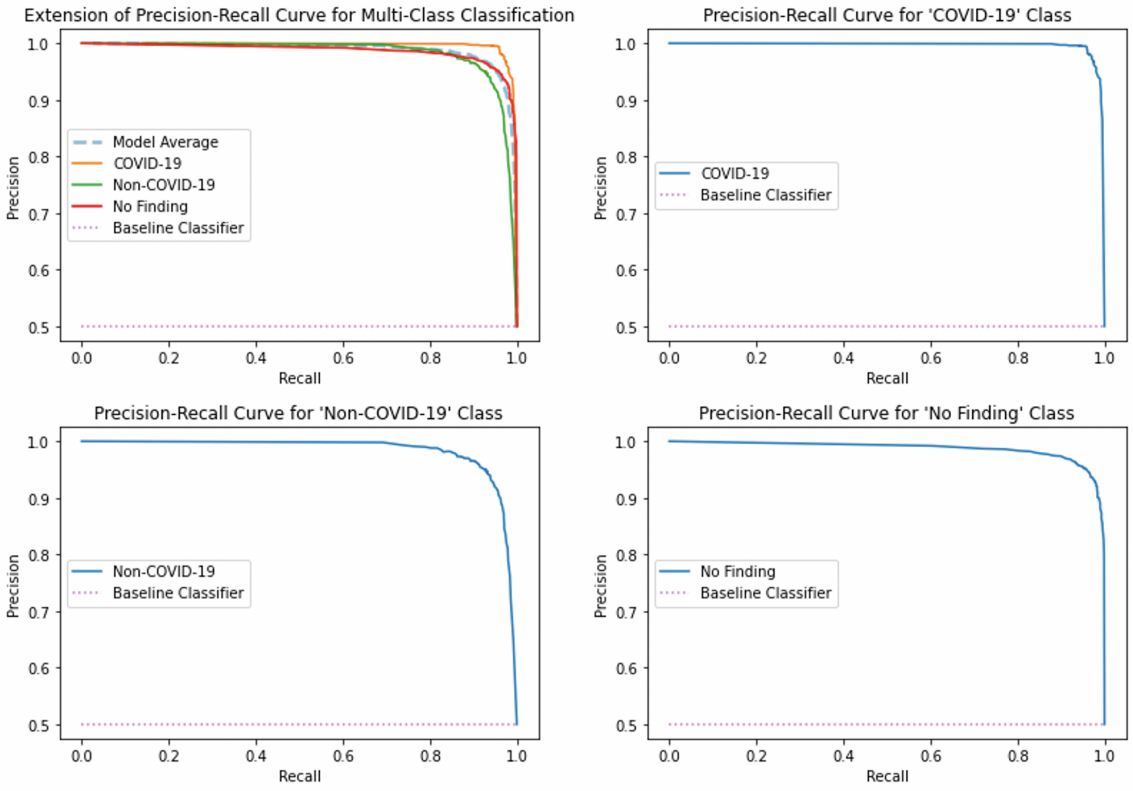

Figure 5. Evaluation of Carebot Covid app using precision-recall curve for each individual class.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2022 doi:10.20944/preprints202203.0288.v1

The precision-recall curve shows the trade-off between precision and recall for different thresholds. (Buckland & Gey 1994) A high area under the curve represents both high recall and high precision, with high precision associated with low False Positive cases and high recall associated with low False Negative cases. (Boyd et al. 2013) Shown baseline classifier is defined as a classifier that cannot distinguish between classes and would predict a random class or a same class in all cases. (Brownlee 2018)

6 Conclusion and future work

In many of the previous studies analyzed, it has been shown that having sufficient clinically annotated data is essential for training more complex CADe applications. (Lebre et al. 2020, Elgendi et al. 2020, Nahiduzzaman et al. 2021) The results of this study showed that the amount of data is also a determining factor for the accuracy of the CNN-based model approach used in the case study. We found that the Carebot Covid app using a deep learning-based model can detect SARS-CoV-2 from CXR images with an accuracy of 0.981, recall of 0.962 and AP of 0.993. In the case of a deep learning-based model, it can be used to detect SARS-CoV-2 from CXR images.

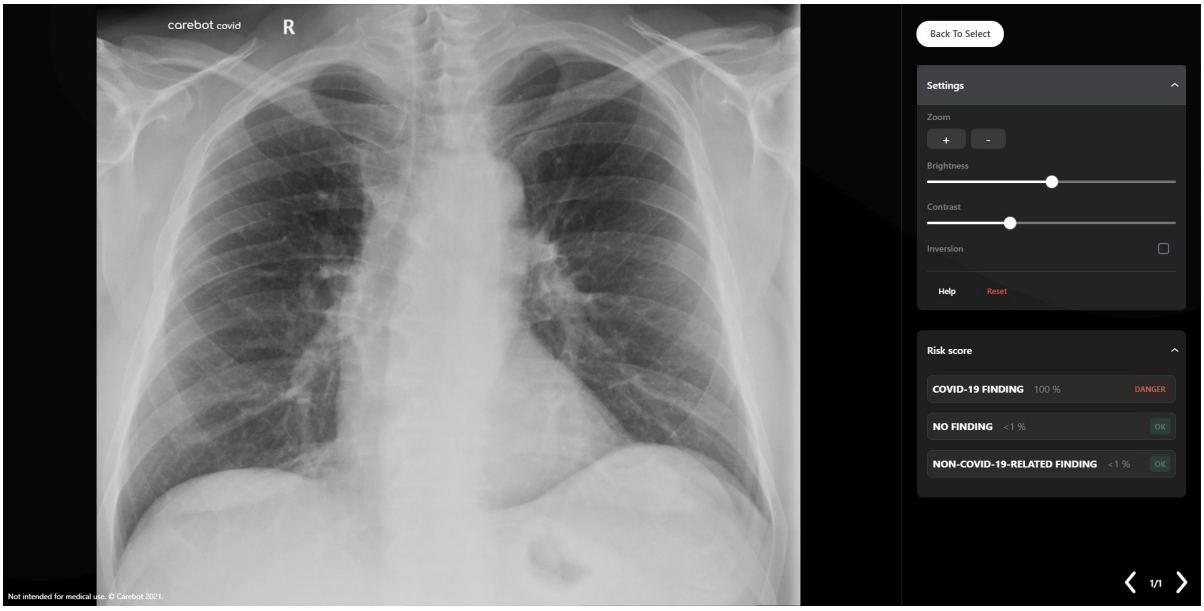

Figure 6. Simplified GUI of Carebot Covid app.

 doi:10.20944/preprints202203.0288.v1

However, data quantity is not the only element that can be central to an CADe system; the machine learning algorithm used and its parameters, as well as the characteristics of the added data, have a significant impact on the final accuracy. (Reiner et al. 2005) As CNN increases its applicability and importance in the medical imaging domain, radiology workflows that enable AI models to access medical data are critically important. (Reiner et al. 2005, Meenatchi Aparna & Shanmugavadivu 2018)

Our design avoids the limitation of usability in the real world by adapting to radiology workflow and adhering to industry norms and standards. Although we did not evaluate the added value of the CADe to the diagnostic performance of radiologists, given the results of previous studies in which radiologists' diagnostic performance improved with deep learning algorithms, we believe that this algorithm can improve radiologists' performance even in large-scale population screening.

Authorship statement

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication.

Ethical procedure

The authors hereby declare that this research article meets all applicable standards with regards to the ethics of experimentation and research integrity. The authors also declare that the text of the article complies with ethical standards, the anonymity of the patients was respected.

 doi:10.20944/preprints202203.0288.v1

References

Albawi, S., Mohammed, T. A. & Al-Zawi, S. (2017), Understanding of a convolutional neural network, *in* '2017 International Conference on Engineering and Technology (ICET)', pp. 1–6.

Apostolopoulos, I. D., Aznaouridis, S. I. & Tzani, M. A. (2020), 'Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases', *Journal of Medical and Biological Engineering* **40**(3), 462–469.

Apostolopoulos, I. D. & Mpesiana, T. A. (2020), 'Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks', *Physical and engineering sciences in medicine* **43**(2), 635–640.

Boyd, K., Eng, K. H. & Page, C. D. (2013), Area under the precision-recall curve: point estimates and confidence intervals, *in* 'Joint European conference on machine learning and knowledge discovery in databases', Springer, pp. 451–466.

Brownlee, J. (2018), 'How to use roc curves and precision-recall curves for classification in python', *Machine learning mastery* **30**.

Buckland, M. & Gey, F. (1994), 'The relationship between recall and precision', *Journal of the American society for information science* **45**(1), 12–19.

Cascella, M., Rajnik, M., Aleem, A., Dulebohn, S. C. & Di Napoli, R. (2022), 'Features, evaluation, and treatment of coronavirus (covid-19)', *Statpearls [internet]* .

Cavallo, J. J. & Forman, H. P. (2020), 'The economic impact of the covid-19 pandemic on radiology practices', *Radiology* **296**(3), E141–E144.

Clunie, D. A., Dennison, D. K., Cram, D., Persons, K. R., Bronkalla, M. D., Primo, H. et al. (2016), 'Technical challenges of enterprise imaging: Himss-siim collaborative white paper', *Journal of digital imaging* **29**(5), 583–614.

Cozzi, D., Cavigli, E., Moroni, C., Smorchkova, O., Zantonelli, G., Pradella, S. & Miele, V. (2021), 'Ground-glass opacity (ggo): A review of the differential diagnosis in the era of covid-19', *Japanese journal of radiology* **39**(8), 721–732.

El Asnaoui, K. & Chawki, Y. (2021), 'Using x-ray images and deep learning for automated detection of coronavirus disease', *Journal of Biomolecular Structure and Dynamics* **39**(10), 3615–3626.

Elgendi, M., Nasir, M. U., Tang, Q., Fletcher, R. R., Howard, N., Menon, C., Ward, R., Parker, W. & Nicolaou, S. (2020), 'The performance of deep neural networks in differentiating chest x-rays of covid-19 patients from other bacterial and viral pneumonias', *Frontiers in Medicine* **7**.
URL: *https://www.frontiersin.org/article/10.3389/fmed.2020.00550*

Elgendi, M., Nasir, M. U., Tang, Q., Smith, D., Grenier, J.-P., Batte, C., Spieler, B., Leslie, W. D., Menon, C., Fletcher, R. R. et al. (2021), 'The effectiveness of image augmentation in deep learning networks for detecting covid-19: A geometric transformation perspective', *Frontiers in Medicine* **8**.

Firmino, M., Angelo, G., Morais, H., Dantas, M. R. & Valentim, R. (2016), 'Computer-aided detection (cade) and diagnosis (cadx) system for lung cancer with likelihood of malignancy', *Biomedical engineering online* **15**(1), 1–17.

Genereaux, B. W., Dennison, D. K., Ho, K., Horn, R., Silver, E. L., O'Donnell, K. & Kahn, C. E. (2018), 'Dicomweb™: Background and application of the web standard for medical imaging', *Journal of Digital Imaging* **31**(3), 321–326.

He, K., Zhang, X., Ren, S. & Sun, J. (2016), Deep residual learning for image recognition, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 770–778.

Hendrycks, D. & Gimpel, K. (2016), 'A baseline for detecting misclassified and out-of-distribution examples in neural networks', *arXiv preprint arXiv:1610.02136* .

Hermann, K., Chen, T. & Kornblith, S. (2020), 'The origins and prevalence of texture bias in convolutional neural networks', *Advances in Neural Information Processing Systems* **33**, 19000–19015.

Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. (2017), Densely connected convolutional networks, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 4700–4708.

Jacobi, A., Chung, M., Bernheim, A. & Eber, C. (2020), 'Portable chest x-ray in coronavirus disease-19 (covid-19): A pictorial review', *Clinical imaging* **64**, 35–42.

Japkowicz, N. & Stephen, S. (2002), 'The class imbalance problem: A systematic study', *Intelligent data analysis* **6**(5), 429–449.

Kagadis, G. C., Kloukinas, C., Moore, K., Philbin, J., Papadimitroulas, P., Alexakos, C., Nagy, P. G., Visvikis, D. & Hendee, W. R. (2013), 'Cloud computing in medical imaging', *Medical physics* **40**(7), 070901.

Khan, A. I., Shah, J. L. & Bhat, M. M. (2020), 'Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images', *Computer Methods and Programs in Biomedicine* **196**, 105581.
URL: *https://www.sciencedirect.com/science/article/pii/S0169260720314140*

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012), Imagenet classification with deep convolutional neural networks, *in* F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger, eds, 'Advances in Neural Information Processing Systems', Vol. 25, Curran Associates, Inc.
URL: *https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf*

Kvak, D. & Kvaková, K. (2021), 'Automatic detection of pneumonia in chest x-rays using lobe deep residual network'.

Le Maitre, A., Fernando, J., Morvan, Y., Mevel, G. & Cordonnier, E. (2014), Comparative performance investigation of dicom c-store and dicom http-based requests, *in* '2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society', pp. 1350–1353.

Le, Y. & Yang, X. S. (2015), Tiny imagenet visual recognition challenge.

Lebre, R., Silva, L. B. & Costa, C. (2020), 'A cloud-ready architecture for shared medical imaging repository', *Journal of Digital Imaging* **33**(6), 1487–1498.

Liu, L., Liu, L., Fu, X., Huang, Q., Zhang, Y., Luo, Q. & Xiong, X. (2015), 'Smartwado: An extensible wado middleware for regional medical image sharing', *Journal of digital imaging* **28**(5), 547–557.

Mahmud, T., Rahman, M. A. & Fattah, S. A. (2020), 'Covxnet: A multi-dilation convolutional neural network for automatic covid-19 and other pneumonia detection from chest x-ray images with transferable multi-receptive feature optimization', *Computers in Biology and Medicine* **122**, 103869.
URL: *https://www.sciencedirect.com/science/article/pii/S0010482520302250*

Martínez Chamorro, E., Díez Tascón, A., Ibáñez Sanz, L., Ossaba Vélez, S. & Borruel Nacenta, S. (2021), 'Radiologic diagnosis of patients with covid-19', *Radiología (English Edition)* **63**(1), 56–73.
URL: *https://www.sciencedirect.com/science/article/pii/S2173510721000033*

Meenatchi Aparna, R. & Shanmugavadivu, P. (2018), A survey of medical imaging, storage and transfer techniques, *in* 'International Conference on ISMAC in Computational Vision and Bio-Engineering', Springer, pp. 17–29.

Mikołajczyk, A. & Grochowski, M. (2018), Data augmentation for improving deep learning in image classification problem, *in* '2018 international interdisciplinary PhD workshop (IIPhDW)', IEEE, pp. 117–122.

Nahiduzzaman, M., Goni, M. O. F., Anower, M. S., Islam, M. R., Ahsan, M., Haider, J., Gurusamy, S., Hassan, R. & Islam, M. R. (2021), 'A novel method for multivariant pneumonia classification based on hybrid cnn-pca based feature extraction using extreme learning machine with cxr images', *IEEE Access* **9**, 147512–147526.

Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O. & Acharya, U. R. (2020), 'Automated detection of covid-19 cases using deep neural networks with x-ray images', *Computers in biology and medicine* **121**, 103792.

Qin, C., Yao, D., Shi, Y. & Song, Z. (2018), 'Computer-aided detection in chest radiography based on artificial intelligence: a survey', *Biomedical engineering online* **17**(1), 1–23.

Reiner, B. I., Siegel, E. L., Hooper, F. J., Siddiqui, K. M., Musk, A., Walker, L. & Chacko, A. (2005), 'Multi-institutional analysis of computed and direct radiography: Part i. technologist productivity', *Radiology* **236**(2), 413–419.

Rothan, H. A. & Byrareddy, S. N. (2020), 'The epidemiology and pathogenesis of coronavirus disease (covid-19) outbreak', *Journal of autoimmunity* **109**, 102433.

Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R. & Sieh, W. (2019), 'Deep learning to improve breast cancer detection on screening mammography', *Scientific reports* **9**(1), 1–12.

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N. & Siddique, R. (2020), 'Covid-19 infection: Emergence, transmission, and characteristics of human coronaviruses', *Journal of advanced research* **24**, 91–98.

Shorten, C. & Khoshgoftaar, T. M. (2019), 'A survey on image data augmentation for deep learning', *Journal of big data* **6**(1), 1–48.

Sokolova, M., Japkowicz, N. & Szpakowicz, S. (2006), Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation, *in* 'Australasian joint conference on artificial intelligence', Springer, pp. 1015–1021.

Speets, A. M., van der Graaf, Y., Hoes, A. W., Kalmijn, S., Sachs, A. P., Rutten, M. J., Gratama, J. W. C., van Swijndregt, A. D. M. & Mali, W. P. (2006), 'Chest radiography in general practice: indications, diagnostic yield and consequences for patient management', *British Journal of General Practice* **56**(529), 574–578.

Tahir, A. M., Chowdhury, M. E., Khandakar, A., Rahman, T., Qiblawey, Y., Khurshid, U., Kiranyaz, S., Ibtehaz, N., Rahman, M. S., Al-Maadeed, S. et al. (2021), 'Covid-19 infection localization and severity grading from chest x-ray images', *Computers in biology and medicine* **139**, 105002.

Wu, Y.-C., Chen, C.-S. & Chan, Y.-J. (2020), 'The outbreak of covid-19: An overview', *Journal of the Chinese medical association* **83**(3), 217.

Yamashita, R., Nishio, M., Do, R. K. G. & Togashi, K. (2018), 'Convolutional neural networks: an overview and application in radiology', *Insights into imaging* **9**(4), 611–629.

Yilmaz, E. & Aslam, J. A. (2006), Estimating average precision with incomplete and imperfect judgments, *in* 'Proceedings of the 15th ACM international conference on Information and knowledge management', pp. 102–111.

Zhang, Y., Liu, M., Hu, S., Shen, Y., Lan, J., Jiang, B., de Bock, G. H., Vliegenthart, R., Chen, X. & Xie, X. (2021), 'Development and multicenter validation of chest x-ray radiography interpretations based on natural language processing', *Communications Medicine* **1**(1), 1–12.