COVID 19 Detection Through Radiograph Images Using Deep Learning Algorithm

Alpas, Michael
Department of Computer Science, UIUC
(309)533-8994
malpas2@illinois.edu

Ahmad, Wasique
Department of Computer Science, UIUC
(414)630-6361
Wasique2@illinois.edu

Singh, Mohit
Department of Computer Science, UIUC
(309)750-7110
Mohits3@illinois.edu

Yadav, Upendra Singh
Department of Computer Science, UIUC
(215)439-4579
Usyadav2@illinois.edu

1. MOTIVATION

The COVID-19 pandemic has led to a dramatic loss of human life worldwide and presents an unprecedented challenge to public health, food systems and the world of work caused by the infection of individuals by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The economic and social disruption caused by the pandemic is devastating tens of millions of people are at risk of falling into extreme poverty, while the number of undernourished people is constantly increasing.

A critical step in the fight against COVID-19 is effective screening of infected patients, such that those infected can receive immediate treatment and care, as well as be isolated to mitigate the spread of the virus. The main screening method used for detecting COVID-19 cases is reverse transcriptase-polymerase chain reaction (RT-PCR) testing, which can detect SARS-CoV-2 RNA from respiratory specimens (collected through a variety of means such as nasopharyngeal or oropharyngeal swabs). While RT-PCR testing is the excellent, it is also very time-consuming, laborious, and complicated manual process that is also in short supply.

An alternative screening method that has also been utilized for COVID-19 screening has been radiography examination, where chest radiography imaging (e.g., chest X-ray (CXR) or computed tomography (CT) imaging) is conducted and analyzed by radiologists to look for visual indicators associated with SARS-CoV-2 viral infection.

The detection of severe acute respiratory syndrome coronavirus 2 (SARS CoV-2), which is responsible for coronavirus disease 2019 (COVID-19), using chest X-ray images has life-saving importance for both patients and doctors.

It was found in early studies that patients present abnormalities in chest radiography images that are characteristic of those infected with COVID-19, with some radiologists suggesting that radiography examination could be used as a primary tool for COVID-19 screening in epidemic areas.

In countries that are unable to purchase laboratory kits for testing, this study becomes even more vital. In this, we aimed to present the use of deep learning for the high-accuracy detection of COVID-19 using chest X-ray images. [13]

1.1 Benefit of using Xray images in detecting COVID19.

Rapid Triaging – CXR (Chest X-Ray) imaging enables rapid triaging of patients suspected of COVID-19 and can be done in parallel of viral testing (which takes time) to help relief the high volumes of patients especially in areas most affected where they have run out of capacity.

Availability of training data - CXR imaging is readily available and accessible in many clinical sites and imaging centers as it is considered standard equipment in most healthcare systems.

Portability - The existence of portable CXR systems means that imaging can be performed within an isolation room, thus significantly reducing the risk of COVID-19 transmission.

Motivated by the urgent need to develop solutions to aid in the fight against the COVID-19 pandemic and inspired by the open source and open access efforts of the research community, and as a deep learning project need, we thought of exploring various deep learning methods specifically a deep convolutional neural network design using pytorch framework to detect COVID-19 cases from CXR images that are open source and available to the general public.

2. LITERATURE

Different Artificial Intelligence tools have produced established and precise results in the applications that use either image-based or other types of data. Apostolopoulos and Mpesiana[4] performed one of the first studies on COVID-19 detection using X-ray images. In their research they evaluated the performance of state-of-the-art convolutional neural network architectures proposed over the recent years for medical image classification.

Experimental results show that their proposed model achieved an overall accuracy of 89.6%, and more importantly the precision and recall rate for COVID-19 cases are 93% and 98.2% for 4-class cases (COVID vs Pneumonia bacterial vs pneumonia viral vs normal). For 3-class classification (COVID vs Pneumonia vs normal), the proposed model produced a classification accuracy of 95%. The preliminary results of this study look promising which can be further improved as more training data becomes available.[3]

Their model achieved promising results on a small, prepared dataset which indicates that given more data, the proposed model can achieve better results with minimum pre-processing of data.

Overall, the proposed model substantially advances the current radiology-based methodology and during COVID-19 pandemic, it can be very helpful tool for clinical practitioners and radiologists to aid them in diagnosis, quantification, and follow-up of COVID-19 cases.

3. DATA

The data used to train the model was chosen from below repository –

https://github.com/ieee8023/covid-chestxray-dataset

This data contains 225 COVID-19 chest X-ray images, which were obtained from Cohen data source[2] and available in above Github link.

This is the very first publicly available COVID-19 X-ray image data set; hence some patient's information is missing in few places.

Dataset contains multiple important metadata fields such as demographic features such as – sex, age, and medical features such as temperature, WBC count, neutrophil count, lymphocyte count, O2 saturation, extubated, intubated, survival and RT_PCR positive.

This dataset contains 131 male patients and 64 female patients and the average age for the group is 58.8±14.9 years.

4. APPROACH

The entire project development cycles can be divided into two stages. Feature Engineering using CNN (Convolutional Neural Network) and Model Training.[5]

- For PyTorch Model Training, we are planning to use ResNet-50, ResNet152 and InceptionV3 or Inception-ResNetV2. [6]
- For building the last layer of binary or multi-level classification, we are thinking to use below classifier algorithms.
    - Logistic Regression.[7]
    - Boosted Random Forest Classification.[8]
    - Decision Tree.[9]
- DL Frameworks [10]
    - PyTorch and PySpark

4.1 CNN

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution.

CNNs are used for image classification and recognition because of its high accuracy. Examples of CNN in computer vision are face recognition, image classification etc.

Convolutional networks are a specialized type of neural networks that use convolution in place of general matrix multiplication in at least one of their layers. It follows a hierarchical model which works on building a network, like a funnel, and finally gives out a fully connected layer where all the neurons are connected to each other and the output is processed.

The main purpose of convolutional layers is to extract features from input images. In this study, different Convolutional layers can be applied to extract different types of features such as edges, texture, colors, and high-lighted patterns from the images.

A convolution layer is simply a feature detection layer. Every convolution layer has a specific number of channels; each channel detects a specific feature in the image. Each feature to detect is often called a kernel or a filter. The kernel is of a fixed size, usually, kernels of size 3 x 3 are used.

In a CNN, the input is a tensor with a shape: (number of inputs) x (input height) x (input width) x (input channels). After passing through a convolutional layer, the image becomes abstracted to a feature map, also called an activation map, with shape: (number of inputs) x (feature map height) x (feature map width) x (feature map channels). A convolutional layer within a CNN generally has the following attributes:

- Convolutional filters/kernels defined by a width and height (hyper-parameters).
- The number of input channels and output channels (hyper-parameters). One layer's input channel must equal the number of output channels (also called depth) of its input.
- Additional hyperparameters of the convolution operation, such as: padding, stride, and dilation.

4.2 Logistic Regression

Logistic regression is a supervised, classification algorithm which is used when the value we are going to predict can be divided into categories. Logistic regression transforms its output using the logistic sigmoid function, calculates its weights and biases to return a probability value which can then be mapped to two or more discrete classes.

As pneumonia and COVID detection is a classification problem, we can use logistic regression to return a probability score between 0 and 1, with 0 representing a normal lung and 1 representing infected lungs.

4.3 Random Forest

Random forest, like its name implies, consists of a large number of individual decision trees that operate as an ensemble. It uses bagging and feature randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree. [11]

A Boosted Random Forest is an algorithm, which consists of two parts; the boosting algorithm: AdaBoost and the Random Forest classifier algorithm[12] which in turn consists of multiple decision trees. A decision tree builds models that are similar to an actual tree. The algorithm divides our data into smaller subsets, simultaneously adding branches to the tree. The outcome is a tree consisting of leaf nodes and decision nodes. A decision node has two or more branches representing the value of each feature (like age, symptom1, etc.) tested and the leaf node holds the result value on the patient's prospective condition (target value).

Multiple classifier decision trees (ensemble of classifiers) eliminate the risk of failure of a single decision tree to correctly predict the target value. Thus, the random forest averages the result provided by multiple trees to provide the final result.

4.4 PyTorch

PyTorch is an open-source machine learning library based on the Torch library that implements a dynamic computational graph, which allows you to change the way your neural network behaves on the fly and capable of performing backward automatic differentiation. It provides two high-level features:

- Tensor computing (like NumPy) with strong acceleration via graphics processing units (GPU)

- Deep neural networks built on a type-based automatic differentiation system.

4.5 PySpark

PySpark is a great language for performing exploratory data analysis at scale, building machine learning pipelines, and creating ETLs for a data platform. It blends the powerful Spark big data processing engine with the Python programming language to provide a data analysis platform that can scale up for nearly any task.

4.6 Performance Criteria

A variety of performance criteria can be used to evaluate the performance of classification models, namely classification accuracy, sensitivity, specificity, precision, recall, and F1 score. Sensitivity and specificity are two suitable criteria that can be used to report model performance. These criteria are also widely used in the medical field.

The cross-validation estimator can be used and can be created a confusion matrix, as shown in table below.

Table 2. Confusion Matrix.

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

- TP (true positive) is a number of anomalies and was identified with the correct diagnosis.
- TN (true negative) is an incorrectly measured number of regular instances.
- FP (false positive) is a Type 1 error, a set of regular instances that are classified as an abnormality diagnosis.
- FN (false negative) is a Type 2 error, a list of abnormalities observed as an ordinary diagnosis.

Accuracy is defined by the rate of correctly classified images.

$$Accuracy = TP + TN / (TP + TN + FP + FN)$$

Precision is used to give relationship between the TP predicted values and FP predicted values.

$$Precision = TP / (TP + FP)$$

Recall is defined as the ratio of the total number of correctly classified positive patients to the total number of positive patients. It should be as high as possible.

$$Recall = TP / (FN + TP)$$

F1 score: It is difficult to compare two models with low precision and high recall, or vice versa. Thus, to render them comparable, we use the F1 score. It allows for the measurement of the Recall and Precision values at the same time. It uses the harmonic mean in place of the arithmetic mean by penalizing the extreme values more.

$$F1\ score = 2 * (precision * recall) / (precision * recall)$$

5. EXPERIMENTAL SETUP

For this experiment, we are planning to use the following technology stack – Docker, Jupyter Notebook and PyTorch. To manage the dependencies, we are planning to leverage containerized environment using Docker. This container will be responsible for hosting a Jupyter notebook environment and for our base image, we will install Python libraries such as PyTorch, NumPy and Pandas. The Docker image will be saved in teams' Docker Hub. The advantage of this kind of set-up, team members will have a ready-to-go consistent working environment so each member can focus on the assigned tasks.

The containerized set-up is a simple solution, however, in the event that a greater processing power is needed due to limitation of team member's individual laptops – we discussed to possibly leverage AWS Elastic Container Service or AWS Fargate for managed serverless compute for containers.

6. TIMELINE

Each team member will be given an opportunity to read several research papers to help us build our ideas and strategy for the project. Individuals will get a chance to do dataset analysis and develop the models we may need, and we will be using the containerized environment. Also, there are several tasks assigned to individuals – Michael will focus on making the container environment which is aiming to be completed by end of 1st week of April. Along the 1st week of April, members will read different research papers and share learned approach we can use for the project. Mohit will start coordinating and drafting the project which should be ready for submission by April 17th.

Wasique and Upendra will start thinking about Deep Learning algorithms after our share-out based on the research papers we read and will be responsible for AWS setup just in case we might need it. Once individual tasks are done, member can join with the analysis and coding. The team is aiming to complete the coding by end of April. Then we prepare to complete the report and presentation by April 5th. Our schedule is relaxed since all of us our working professionals and have families to spend time.

7. REFERENCES

[1] Roberts, M., Driggs, D., Thorpe, M. *et al.* Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* **3,** 199–217 (2021). (link: https://doi.org/10.1038/s42256-021-00307-0)

[2] Cohen, J. P. COVID-19 Image Data Collection. ArXiv. 2020, arXiv:2003.11597. https://doi.org/10.1177/2472630320958376

[3] Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed. 2020 Nov;196:105581. doi: 10.1016/j.cmpb.2020.105581. Epub 2020 Jun 5. PMID: 32534344; PMCID: PMC7274128.

"https://pubmed.ncbi.nlm.nih.gov/32534344/"

[4] Apostolopoulos, I.D., Mpesiana, T.A. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43, 635–640 (2020).

https://doi.org/10.1007/s13246-020-00865-4

[5] Z. Zahisham, C. P. Lee and K. M. Lim, "Food Recognition with ResNet-50," 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2020, pp. 1-5, doi: 10.1109/IICAIET49801.2020.9257825. (link: link: https://doi.org/10.1016/j.procs.2018.05.069)

[6] https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035

[7] Peng, CY.J., So, TS.H., Stage, F.K. et al. The Use and Interpretation of Logistic Regression in Higher Education Journals: 1988–1999. Research in Higher Education 43, 259–293 (2002). https://doi.org/10.1023/A:1014858517172

[8] Mariana Belgiu, Lucian Drăguţ,Random forest in remote sensing: A review of applications and future directions, ISPRS Journal of Photogrammetry and Remote Sensing, https://doi.org/10.1016/j.isprsjprs.2016.01.011. (https://www.sciencedirect.com/science/article/pii/S0924271616000265)

[9] Yoo, S. H., Geng, H., Chiu, T. L., Yu, S. K., Cho, D. C., Heo, J., Choi, M. S., Choi, I. H., Cung Van, C., Nhung, N. V., Min, B. J., & Lee, H. (2020). Deep Learning-Based Decision-Tree Classifier for COVID-19 Diagnosis From Chest X-ray Imaging. Frontiers in medicine, 7, 427. https://doi.org/10.3389/fmed.2020.00427

[10] https://blog.usejournal.com/deep-learning-with-pytorch-zero-to-gans-week-1-592397473811

[11] Md. Zahangir Alam, M. Saifur Rahman, M. Sohel Rahman, A Random Forest based predictor for medical data classification using feature ranking, https://doi.org/10.1016/j.imu.2019.100180.(https://www.sciencedirect.com/science/article/pii/S235291481930019X)

[12] Iwendi Celestine, Bashir Ali Kashif, Peshkar Atharva, Sujatha R., Chatterjee Jyotir Moy, Pasupuleti Swetha, Mishra Rishita, Pillai Sofia, Jo Ohyun TITLE=COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm (link: https://www.frontiersin.org/article/10.3389/fpubh.2020.00357

[13] Khanday, A.M.U.D., Rabani, S.T., Khan, Q.R. et al. Machine learning based approaches for detecting COVID-19 using clinical text data. Int. j. inf. tecnol. 12, 731–739 (2020). https://doi.org/10.1007/s41870-020-00495-9