SARS COVID-19 Detection through AI

ABSTRACT

Early diagnosis of COVID-19 may aid in the development of effective treatment options and the implementation of measures to contain the condition. Using pictures from X-ray, ultrasound, and CT scans, the three most used medical imaging modalities, we show how transfer learning from deep learning models may be utilized to conduct COVID-19 detection. The goal is to use sophisticated deep learning image categorization algorithms to act as a second set of eyes for overworked medical personnel. By preliminary comparison research, we determine a viable convolutional neural network (CNN) model.

I. INTRODUCTION

Almost 18.35 million people have been infected and over 6,96,147 have died as a result of the current COVID-19 epidemic. The best way to control this epidemic is to find and isolate patients as soon as possible. Our research hopes to lay a theoretical groundwork.

‘‘transfer learning’’ whereby a deep learning network is pre-weighted with the results of a previous training cycle from a different domain. This technique is commonly used as a basis for initializing deep learning models.

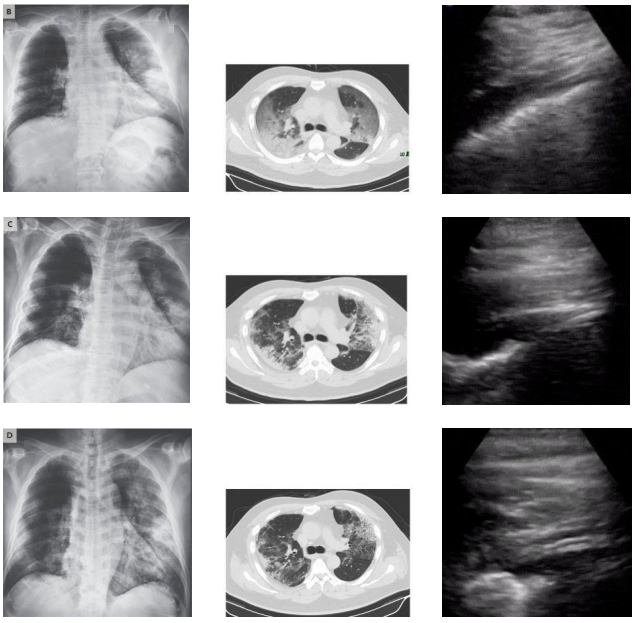
The study willdemonstrate how transfer learning can be used for COVID- 19 detection for three commonly used imaging modes X- Ray, Ultrasound, and CT scan. This could assist practitionersand researchers in developing a supporting tool for highly constrained health professionals in determining the course oftreatment. The study further demonstrates a pre-processing pipeline for improving the image quality, for deep learning- based predictions. An initial testing is also conducted to understand the suitability of various popular deep learning models for the limited available dataset in order to select a model for the proposed image classification demonstrations on multiple image modes.

X-Ray imaging is relatively cost effective and commonly utilized for lung infection detection and is useful for COVID-19 detection as well [15]. Medical observations were made by one of the co-authors of this research who is also a medical professional, as well as by treating doctors ofthe COVID-19 dataset [16] patients.

transfer learning framework to support COVID-19 detection with the use of image classification using deep learning models for multiple imaging modes including X-Ray, Ultrasound, and CT scan. The acquisition of a sufficiently large, publicly available corpus of medical image sample data for fully training deep learning models is challenging for novel medical conditions such as COVID-19 since collection and labelling of images requires significant time and resources to compile. An alternative method of training deep learning models is by one of the co-authors of this research who is also a medical professional, as well as by treating doctors ofthe COVID-19 dataset [16] patients.

The common features observed in the X-Ray images of patients with COVID-19 are patchy infiltrates or opacities that bear similarities to other viral pneumonia features. X-Ray images do not show any abnormalities in the early stages of COVID-19.

However,as the disease progresses, COVID-19 gradually manifests asa typical unilateral patchy infiltration involving mid zone andupper or lower zone of the lungs, occasionally consolidation.



Ultrasound imaging has also been recommended as a tool for COVID-19 lung condition assessment since it can be used at bedside with minimal infection spreading risks and has excellent ability to detect lung conditions related to COVID-19 Progression of COVID-19 infection is evident as B-lines aeration in early stages of consolidation in critical stages.

Fig. 1 shows the progression of evidence for the patient in the COVID-19 datasets for X-Ray, CT and Ultrasound imaging.

Computer vision diagnostic tools for COVID-19 from multiple imaging modes such as X-Ray, Ultrasound, and CT would provide an automated ‘‘second reading’’ to clinicians, assisting in the diagnosis and criticality assessment of COVID-19 patients to assist in better decision making in the global fight against the disease. COVID-19 often results in pneumonia, and for radiologists and practitioners differentiating between the COVID-19 pneumonia and other types of pneumonia (viral and bacterial) solely based on diagnostic images could be challenging.

Deep learning artificial neural networks, and the Convolutional Neural Networks (CNNs) have proven to be highly effective in a vast range of medical image classification applications. In this study, we present three key contributions. Primarily, we demonstrate how transfer learning capabilities of off the shelf deep learning models can be utilized to perform classifications in two distinct scenarios for three imaging modes X-Ray, Ultrasound, and CT scan:

Deep learning artificial neural networks, and the Convolutional Neural Networks (CNNs) have proven to be highly effective in a vast range of medical image classification applications In this study, we present three key contributions. Primarily, we demonstrate how transfer learn-ing capabilities of off the shelf deep learning models can be utilized to perform classifications in two distinct scenarios for three imaging modes X-Ray, Ultrasound, and CT scan:

1) Identifying the pneumonia (both COVID-19 and other types) affected lung against the normal lung.
2) Identifying COVID-19 affected lung from non COVID- 19 pneumonia affected lung.

Exhaustive performance comparison of these methods, rather we wanted to select the most suitable one for our multi-modal image classification, which performs decently with minimal tuning. The source X-Ray, Ultrasound, and CT image samples, especially those from the COVID-19 data sets, have been harvested from multiple sources and are of inconsistent quality. In our final contribution, we have implemented a pre-processing pipeline to reduce unwanted signal noise such as non-lung area visible in X-Rays, and thereby reduce the impact of sampling bias on this comparison. Through this pre-processing pipeline, we minimize the image quality imbalances in the image samples. This wouldallow models to train on lung features only thus having a greater chance of learning disease features and ignoring othernoise features.
The study would provide timely model selection guidelines to the practitioners who often are resortedto utilize certain mode of imaging due to time and resource scarcity.

II. Although CT scans are much higher contrast/resolution compared to X-Ray factors such as low dose and improper use of image enhancement can lead to poor quality images. Anumber of researchers have noted that histogram equalization techniques, particularly adaptive histogram equalization canimprove the contrast of CT images. A combination of histogram normalization, gamma correction and contrast lim- ited adaptive histogram equalization has been shown to objec- tively improve the quality of poor contrast CT images.

III. Ultrasound images tend to be noisy due to the relatively low penetration of soundwaves into organic tissue comparedto X-Rays. This limitation has led a number of researchers todevelop methods to improve the quality of ultrasound imagesby various means including noise filtering, wavelet transfor-mation and deconvolution . Contrast Limited Adaptive Histogram Equalization (CLAHE) has been used as part of apre-processing pipeline to enhance the quality of ultrasound images.

Although CT scans are much higher contrast/resolution compared to X-Ray factors such as low dose and improper use of image enhancement can lead to poor quality images. Anumber of researchers have noted that histogram equalization techniques, particularly adaptive histogram equalization canimprove the contrast of CT images.

Ultrasound images tend to be noisy due to the relatively low penetration of sound waves into organic tissue comparedto X-Rays. This limitation has led a number of researchers todevelop methods to improve the quality of ultrasound imagesby various means including noise filtering, wavelet transformation and deconvolution.Contrast Limited Adaptive Histogram Equalization (CLAHE) has been used as part of apre-processing pipeline to enhance the quality of ultrasound images.

IV.

Computer aided detection and diagnosis of pulmonary pathologies from X-Ray images is a field of research that started in the 1960s and steadily progressed in the following

CT scans also use X-Rays as a radiation source, how- ever, they provide much higher image resolution and contrast compared to standard X-Ray images because of a much more focused X-Ray beam used to produce cross-sectional images of the patient . CT is generally considered as thebest imaging modality for lung parenchyma and is widely accepted by clinicians as the ‘‘gold standard’’.A large corpus of research exists relating to the use of machine learn-ing to improve the efficiency and accuracy of lung cancer diagnosis – largely driven by extensive CT based lung can- cer screening programs in many parts of the world. Several researches have achieved incredibly accurate results using Recently a deep learning system built by Google achieved state-of-the-art performance using patients’ current and priorCT volumes to predict the risk of lung cancer.

This system outperformed human radiologists where prior CT scans werenot available, and equaled human radiologist performance where historical CT scans were available. Although X- Ray is the current reference diagnosis for pneumonia, some studies point out that CT generally outperforms X-Ray asa diagnostic tool for pneumonia, albeit at higher cost and convenience .

V. DATASET DEVELOPMENT

DATA SOURCING

Large numbers of X-Ray, CT and Ultrasound images are available from several publicly accessible datasets. With the emergence of COVID-19 being very recent none of these large repositories contain any COVID-19 labelled data.

COVID-19 chest X-Rays were obtained from the publicly accessible COVID-19 Image Data Collection.This collection has been sourced from websites and pdf format publications. Unsurprisingly, the images from this collection are of variable size and quality. Image contrast levels, brightness and subject positioning are all highly variable within this dataset..

MODEL DEVELOPMENT

The experiment pipeline is shown in Fig. 4. Unprocessed images are read with directory names used as class labels. N-CLAHE is then applied to normalize images and highlight the finer details for the attention of the machine learning classifiers. Images are then resized to the classifier default size, for example 224 $\times$ 224 pixels for VGG16/19 and 299$\times$ 299 pixels for InceptionV3. Following image resizing, data augmentation is applied to increase the number and variation of images provided to the classifier. Augmentations applied include horizontal flip, rotation, width shift, and height shift. Vertical flip was not applied since X-Ray images are not vertically symmetrical, and the resulting flipped image would not resemble a real chest X-Ray. Finally, the augmented images are utilized by the machine learning classifier using an 80:20 Train/Test split.

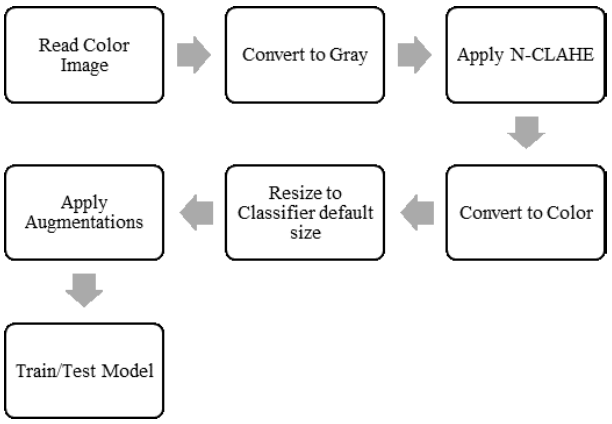


FIGURE 2. Experiment pipeline for preprocessing and classification

The selection of a dataset for Normal and Pneumonia condition X-Rays posed a dilemma since a highly curated data set is not comparable to the available COVID-19 chest X-Ray dataset. Our early tests against one such dataset gave an unrealistically high classification accuracy for the quality of the data under test. We found that the National Institute of Health (NIH) Chest X-Ray [67] dataset – provided images are of a similar size, [68] quality and aspect ratio to the typical images in the COVID-19 dataset with dimensions being uniformly 1024 $\times$ 1024 pixels in a portrait orientation.

CT scans for COVID-19 and non COVID-19 were obtained from the publicly accessible COVID-CT Dataset [66]. This dataset has been sourced by extracting CT slice images showing the COVID-19 pathology from preprint papers. Once again, the images from this collection are of variable size and

TABLE 3. Model performance summary.

Model	Image mode	Training Notes	Testing F1 (avg)
VGG16	X-Ray	Converged well. Overfitting evident at 30 epochs.	0.79
	Ultrasound	Converged well. Overfitting not evident at 100 epochs.	0.99
	CT	Converged. Overfitting evident from 5 epochs.	0.79
VGG19	X-Ray	Converged well. Overfitting not evident at 100 epochs.	0.87
	Ultrasound	Converged well. Overfitting not evident at 100 epochs.	0.99
	CT	Converged. Overfitting evident after 20 epochs.	0.78
Xception	X-Ray	Converged poorly. Overfitting evident after 10 epochs	0.76
	Ultrasound	Converged. Overfitting evident after 20 epochs.	0.82
	CT	Did not converge. Overfitting evident immediately	0.70
InceptionResNet	X-Ray	Converged poorly. Overfitting evident after 10 epochs.	0.73
	Ultrasound	Converged. Overfitting evident after 5 epochs.	0.89
	CT	Did not converge. Overfitting evident after 1st epoch.	0.63
InceptionV3	X-Ray	Converged poorly. Overfitting evident after 10 epochs.	0.75
	Ultrasound	Converged. Overfitting evident after 10 epochs.	0.87
	CT	Did not converge. Overfitting evident after 2 epochs.	0.71
NASNetLarge	X-Ray	Did not converge. Overfitting evident after 10 epochs.	0.64
	X-Ray		0.80
	Ultrasound		0.64
	CT	Converged poorly. Overfitting evident immediately.	
DenseNet121	X-Ray	Converged. Overfitting evident after 8 epochs.	0.74
	Ultrasound	Converged. Overfitting evident after 20 epochs.	0.66
	CT	Converged well. Overfitting evident immediately.	0.75
ResNet50V2		Converged poorly. Overfitting evident immediately.	0.75
	Ultrasound		0.93
	CT		0.66

COMPUTING EQUIPMENT

All the experiments were performed on the University of Technology Sydney Interactive HPC environment under an Anaconda 3 software environment. Experiments were programmed using the Keras APIs with a TensorFlow 2 backend.

The server used was specified as an Intel Xeon Gold 6150 2.7GHz 18 cores (16 cores enabled) with 24.75MB L3 Cache (Max Turbo Freq. 3.7GHz, Min 3.4GHz). The server had 360GB RAM (Six Channel). This server hosted a NVIDIA Quadro P5000 GPU (2,560 Cores, 16GB Memory).

EXPERIMENT SETUP

VGG19 model was tuned for each image mode and each individual experiment to achieve the best possible results for the collated datasets. Learning rates were varied between 10^{-3} to 10^{-6} with order-of-magnitude increments. The batch sizes between 2 to 16 were applied. The hidden layer was varied between 4 and 96 nodes. Dropout rates of 0.1 and 0.2 were applied. These hyperparameter ranges generated an output head architecture as shown in Fig. 5.

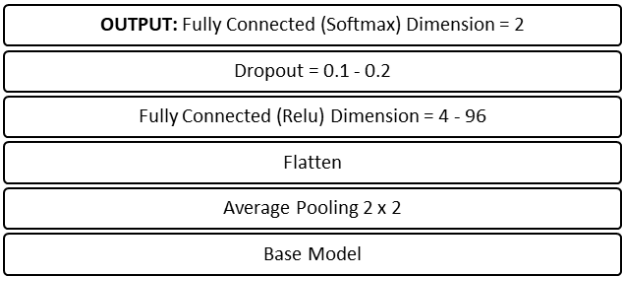


FIGURE 3. Head architecture of the proposed models.

RESULTS AND DISCUSSIONS

With the selected VGG19 model for each experiment listed in Table 4, we first conducted the extensive performance tuning by adjusting multiple parameters including learning rate, batch size, node size and drop rate. The effects of learning rate, batch size and hidden layer size hyperparameter selection on the accuracy metric of experiment 1A is shown in Fig. 6. We noted that dropout rate had only a minimal effect on the model accuracy except at the highest learning rate of 10^{-3} and lowest learning rate of 10^{-6} where a dropout rate of 0.2 proved to be more stable than a dropout rate of 0.1. For learning rates for 10^{-3} and 10^{-4} the dropout rate selection has no discernable effect on model accuracy.

Learning rate, batch size and hidden layer size all affected model accuracy. The first observation is that learning rates

TABLE 1. Datasets used for experiments.

Exp ID	Image Mode	Experiment	Dataset
1A	X-Ray	Normal vs (COVID-19 and Pneumonia)	(400 x Normal) vs (190 x Pneumonia 139 x COVID-19)
1B	Ultrasound	Normal vs (COVID-19 and Pneumonia)	226 x Normal vs (235 x COVID-19 220 x Pneumonia)
2A	X-Ray	COVID-19 vs Pneumonia	139 x COVID-19 vs 190 x Pneumonia
2B	Ultrasound	COVID-19 vs Pneumonia	235 x COVID-19 vs 220 x Pneumonia
3A	CT	COVID-19 vs Non COVID	349 x COVID-19 vs 397 x Non COVID

of 10^{-4} and 10^{-5} provided higher model accuracy with 10^{-5} achieving more consistent results. There is a tendency for accuracy to improve with batch size increase at learning rates of 10^{-3} and 10^{-6} but at learning rates of 10^{-4} and 10^{-5} this tendency is not apparent. Finally, there is also a trend towards higher accuracy with a larger hidden layer size that is most noticeable at 10^{-3}. Taking a learning rate of 10^{-5} as achieving consistent high accuracy, we can then suggest from this analysis that a hidden layer size ranging from 64 to 96 and batch size of 4 could generally be expected to provide the most accurate results for this experiment. Through similar analysis for each experiment in Table 4, we have identified the best parameter settings for each experiment as shown in Table 5.

The results of the five experiments are listed in Table 5. For experiments classifying COVID-19 and Pneumonia vs Normal (1A and 2A) we found that the Ultrasound mode provided the best results with a sensitivity of 97% and positive predictive value of 99% compared to X-Ray with 83% and 85% respectively. For experiments classifying COVID-19 vs Pneumonia (1B and 2B) we again found that the Ultrasound mode provided the best results with a sensitivity of 100% and a positive predictive value of 100% compared to X-Ray with sensitivity of 86% and positive predictive value of 86%. The CT imaging mode was found to have a sensitivity of 83% and positive predictive value of 79% in classifying COVID-19 vs non COVID-19 scans. All experiments resulted in F1 scores exceeding 80% which is a good result given the relatively small and variable quality data corpus available.

The learning curves for each experiment are shownin Fig. 7. The training curves for both Ultrasound experi- ments (2A and 2B) are close to ideal. The training curvesfor the X-Ray experiments (1A and 2A) are also very good, although the curve for experiment 1B does show some signs

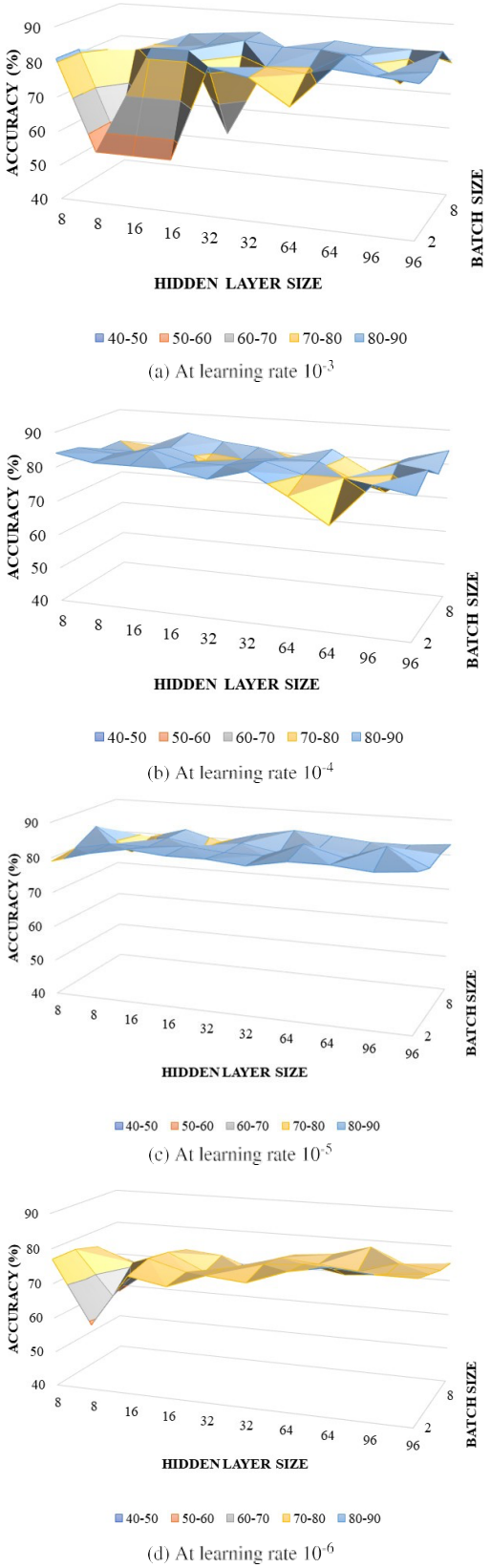


FIGURE 6. Model sensitivity to hyperparameters for experiment 1A.

CONCLUSION

We have demonstrated that with current limited and Challenging COVID-19 datasets, VGG19 model could be With very little data duration, we achieved considerable classification results using VGG19 from all imaging modes. Perhaps the most interesting observation is that the pre-trained models tuned very effectively for the Ultrasound image samples, which to the untrained eye appeared noisy interpret.

Both training curves and confu- sion matrix for both Ultrasound experiments are close to ideal. VGG19 also trained well against the X-Ray image corpus however, thresholding. we found that the proportion of false negatives was concerning but notunexpected given data quality challenges. Our finding that experiment 1A/2A yielded lower F1 scores and higher false negatives than experiments 1B/2B was unexpected since themanifestation of COVID-19 .

This may indicate that despite our attempts to remove sampling bias using N-CLAHE pre-processing there may still be systematic differences in the COVID-19 image data sets that leads the VGG19 classifier to more easily distin- guish the COVID-19 images from the pneumonia images. A future research direction could be to isolate the lung field by segmentation for all image samples in order to remove noise and further reduce sampling bias. Our lower results against the CT image corpus were not surprising since the CT image slices available were not from a uniform patient

REFERENCES

[1] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, ‘‘Correlation of chest CT and RT-PCR testing for coro- navirus disease 2019 (COVID-19) in China: A report of 1014cases,’’ *Radiology*, vol. 296, no. 2, pp. E32–E40, Aug. 2020, doi: 10.1148/radiol.2020200642.

[2] M. Hosseiny, S. Kooraki, A. Gholamrezanezhad, S. Reddy, and L. Myers, ‘‘Radiology perspective of coronavirus disease 2019 (COVID-19): Lessons from severe acute respiratory syndrome and middle east respiratory syndrome,’’ *Amer. J. Roentgenology*, vol. 214, no. 5,pp. 1078–1082, May 2020, doi: 10.2214/AJR.20.22969.

[3] A. Ulhaq, A. Khan, D. Gomes, and M. Paul, ‘‘Computer vision for COVID-19 control: A survey,’’ 2020, *arXiv:2004.09420*. [Online]. Avail- able: http://arxiv.org/abs/2004.09420

[4] J. Born, G. Brändle, M. Cossio, M. Disdier, J. Goulet, J. Roulin, and N. Wiedemann, ‘‘POCOVID-net: Automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS),’’ 2020, *arXiv:2004.12084*. [Online]. Available: http://arxiv.org/abs/2004.12084

[5] D. J. Bell. (2020). *COVID-19 Radiopaedia*. [Online]. Available: https://radiopaedia.org/articles/covid-19-4?lang=us