

# X-Ray Image based COVID-19 Detection using Pre-trained Deep Learning Models

Michael J Horry<sup>1,3</sup>, Subrata Chakraborty<sup>1\*</sup>, Manoranjan Paul<sup>2</sup>, Anwaar Ulhaq<sup>2</sup>, Biswajeet Pradhan<sup>1</sup>, Manash Saha<sup>4</sup>, Nagesh Shukla<sup>1</sup>

<sup>1</sup>School of Information, Systems & Modelling, Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW 2007, Australia.

<sup>2</sup>Machine Vision and Digital Health (MaViDH), School of Computing and Mathematics, Charles Sturt University, Bathurst, NSW 2795, Australia.

<sup>3</sup>IBM Australia Limited.

<sup>4</sup>Manning Rural Referral Hospital, Taree NSW 2430, Australia

\*Email: Subrata.Chakraborty@uts.edu.au

**Abstract**—Detecting COVID-19 early may help in devising an appropriate treatment plan and disease containment decisions. In this study, we demonstrate how pre-trained deep learning models can be adopted to perform COVID-19 detection using X-Ray images. The aim is to provide over-stressed medical professionals a second pair of eyes through intelligent image classification models. We highlight the challenges (including dataset size and quality) in utilising current publicly available COVID-19 datasets for developing useful deep learning models. We propose a semi-automated image pre-processing model to create a trustworthy image dataset for developing and testing deep learning models. The new approach is aimed to reduce unwanted noise from X-Ray images so that deep learning models can focus on detecting diseases with specific features from them. Next, we devise a deep learning experimental framework, where we utilise the processed dataset to perform comparative testing for several popular and widely available deep learning model families such as VGG, Inception, Xception, and Resnet. The experimental results highlight the suitability of these models for current available dataset and indicates that models with simpler networks such as VGG19 performs relatively better with up to 83% precision. This will provide a solid pathway for researchers and practitioners to develop improved models in the future.

**Keywords**—COVID-19 detection, CNN models, X-Ray based detection, image processing, model comparison

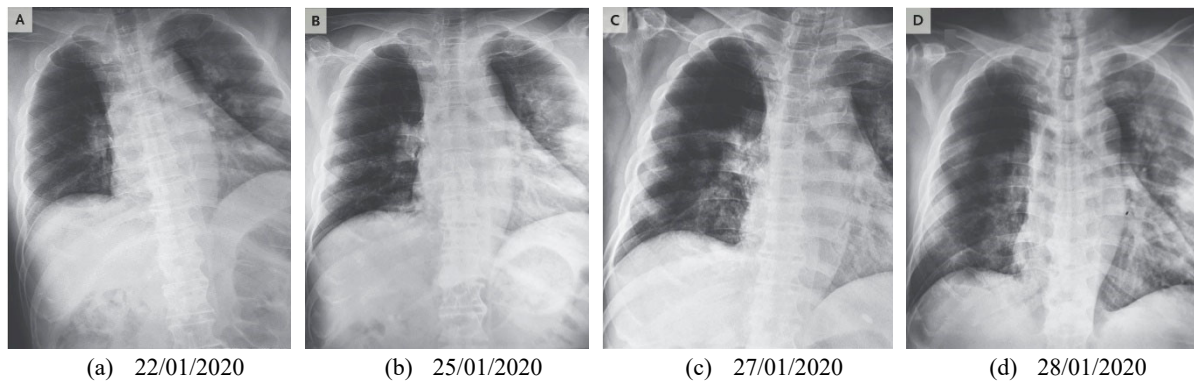
## I. INTRODUCTION

The current COVID-19 pandemic has impacted the world with over 2.3 million infections and close to 160,000 deaths so far (as of 19<sup>th</sup> April 2020). Early identifying, isolation and care for patients is a key strategy for a better management of this pandemic [1]. Our study aims to provide a conceptual framework to support COVID-19 detection with the use of image classification using deep learning models. The study will demonstrate how deep learning can be used for COVID-19 detection, which could be used as a supporting tool for highly constrained health professionals in determining the course of treatment. The study further develops a pre-processing pipeline for improving the image quality, for deep learning-based predictions. We also provide a comparative study between several pre-trained deep learning models which will provide insights for further system development in this application domain.

Fast, accessible, affordable and reliable identification of COVID-19 pathology in an individual is key to slowing the transmission of COVID-19 infection. Currently, reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) tests are the gold standard for diagnosing COVID-19 [2]. During this test small amounts of viral RNA are extracted from a nasal swab, amplified, and quantified with virus detection indicated visually

using a fluorescent dye. Unfortunately, the RT-qPCR test is manual and time-consuming with results taking up to two days. Some studies have also shown false positive Polymerase Chain Reaction PCR testing [3]. Other testing approaches include imaging technology-based approaches including computed tomography (CT) imaging [4] and X-Ray imaging based [5][43]. The CT scan-based COVID-19 detection is time consuming and manual with the requirement of expert involvements. Several automated approaches have been proposed recently [4, 6, 7]. Both the PCR tests and CT scans are comparatively costly [8, 9] and with high demand many countries are forced to perform selective testing for only high risk population.

X-Ray imaging is relatively cost effective and commonly done for lung infection detection and is useful for COVID-19 detection as well [10]. Medical observations were made by one of the co-author of this research (Dr. Saha) who is also a medical professional, as well as by treating doctors of the COVID-19 dataset [11] patients. The common features observed in the X-Ray images of patients with COVID-19 is a patchy infiltration 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 as a typical unilateral patchy infiltration involving mid zone and upper or lower zone of the lungs, occasionally with evidence of a consolidation. Figure 1 shows the progression of evidence for a patient in the COVID-19 dataset.



**Figure 1:** COVID 19 progression over several days for a sample patient.

Computer vision diagnostic tools for COVID-19 from X-Ray images would provide an automated “second reading” to clinicians, assisting in the diagnosis of COVID-19 and informing better decision making in the global fight against the disease. COVID-19 often results in pneumonia, and for radiologists and practitioners differentiating between COVID-19 pneumonia and other types of pneumonia (viral and bacterial) solely on the basis of X-Ray images could be challenging [12].

Deep learning artificial neural networks, and in particular the *Convolutional Neural Networks* (CNNs) have proven to be highly effective in a vast range of medical image classification applications [13, 14]. In this study, we comparatively tested five (5) common off-the-shelf pre-trained CNNs namely VGG16/VGG19 [15], Resnet50 [16], Inception V3 [17] and Xception [18] to determine which CNN implementation is the most effective within the limitations of the publicly available COVID-19 X-Ray image samples [11]. The key goals of our experiments with these models is to find the most suitable deep learning model applicable for the available data for COVID-19 detection from X-Ray images. We aim to apply the models with two distinct scenarios:

- (a) Identifying pneumonia (both COVID-19 and other types) affected lung against normal lung.
- (b) Identifying COVID-19 affected lung from non COVID-19 pneumonia affected lung.

Noting that the source X-Ray image samples, especially those from the COVID-19 data set, have been harvested from multiple sources and are of inconsistent quality 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 only identify lung area and remove rest of the X-Ray image. This would allow models to train on lung features only thus having a greater chance of learning disease features and ignoring other noise features.

In the following sections we first present a brief review of recent scholarly works related to this study, followed by a discussion on the available datasets we used and related challenges. We then present the dataset generation process along with our proposed pre-processing pipeline for data quality balancing. We then present the deep learning model selection process along with comparative results and discussions.

## II. RELATED WORK

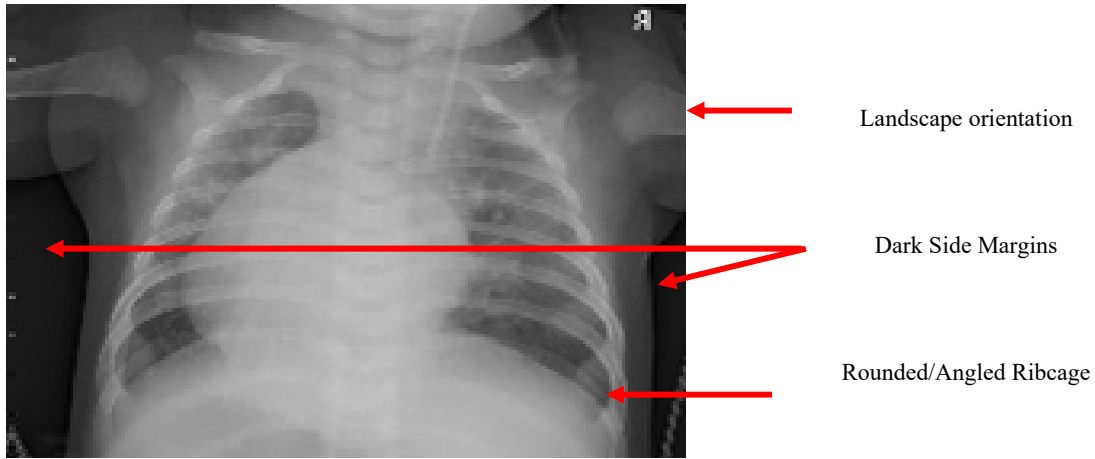
Computer aided detection and diagnosis of pulmonary pathologies from X-Ray images is a field of research that started in the 1960s and steadily progressed in the following decades with papers describing highly accurate diagnosis of a range of conditions including osteoporosis [19], breast cancer [20], and cardiac disease [21].

Difficulties in distinguishing soft tissue caused by poor contrast in X-Ray images have led some researchers to implement contrast enhancement [22] as a pre-processing step in X-Ray based diagnosis. In addition, lung segmentation of X-Ray images is an important step in the identification of lung nodules and various segmentation approaches are proposed in the literature based on linear filtering/thresholding, rolling ball filters and more recently CNNs [23].

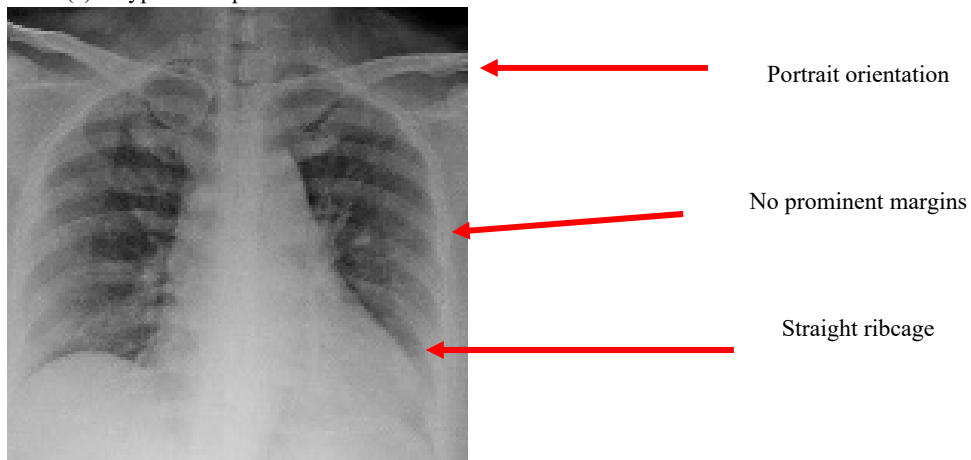
In our literature review we noted a small number of very recent studies that have used deep learning systems for COVID-19 screening and diagnosis. A custom-built 18-layer residual network pre-trained on the ImageNet weights against COVID-19 (100 images) and Pneumonia (1431 images) X-Ray image datasets [24]. A range of deep learning frameworks coined as COVIDX-Net trained on a small data set of 25 confirmed COVID-19 cases [25]. A custom curated dataset of COVID-19, viral pneumonia and normal X-Ray images [26]. A custom residual CNN that was highly effective in distinguishing between COVID-19, Pneumonia and normal condition X-RAY images [27]. These studies used the COVID-19 dataset [11] for the COVID-19 X-Ray samples and the RSNA dataset [28] was used to get pneumonia and normal X-Ray samples. Although some of these studies showed promising results, they have not addressed the variable quality of the COVID-19 data set or minimisation of sampling bias. We also noted that the experimental datasets used in these studies are quite unbalanced. Class imbalance is a common problem in medical image data sets that can lead to over-classification of the majority class at the expense of under-classification of the minority class [29].

The COVID-19 dataset [11] is probably the only publicly available source for some COVID-19 X-Ray images at present. The quality of these images is highly inconsistent since these images were shared from different parts of the world and are often taken in non-ideal conditions and vastly different equipment. On the other hand, the RSNA dataset [28] showed more uniformity and appeared to be highly curated. The quality variation between these two datasets are often readily apparent to the untrained eye as shown in Figure 2. The RSNA dataset also contains a large number of patients who are juvenile. The COVID-19 dataset however contains mostly adult patients. These discrepancies between these datasets are mostly ignored by currently available studies. This may have impacted the performance of the models they developed, since there is a good chance that the deep learning models learned the quality discrepancies between the two datasets rather than characteristics of the diseases.

We believe the significant quality variations between data from different classes need to be balanced for deep learning models to learn actual disease related variations. Therefore, our study stresses the importance of sampling bias/signal noise removal from the X-Ray image datasets prior to using them for model development and classification in order to obtain meaningful and trustworthy classification results. We present the dataset creation process in the next section.



(a) Typical sample from RSNA dataset



(b) Typical Sample from the COVID19 dataset

**Figure 2:** Visual comparison of typical samples from the RSNA and COVID-19 datasets.

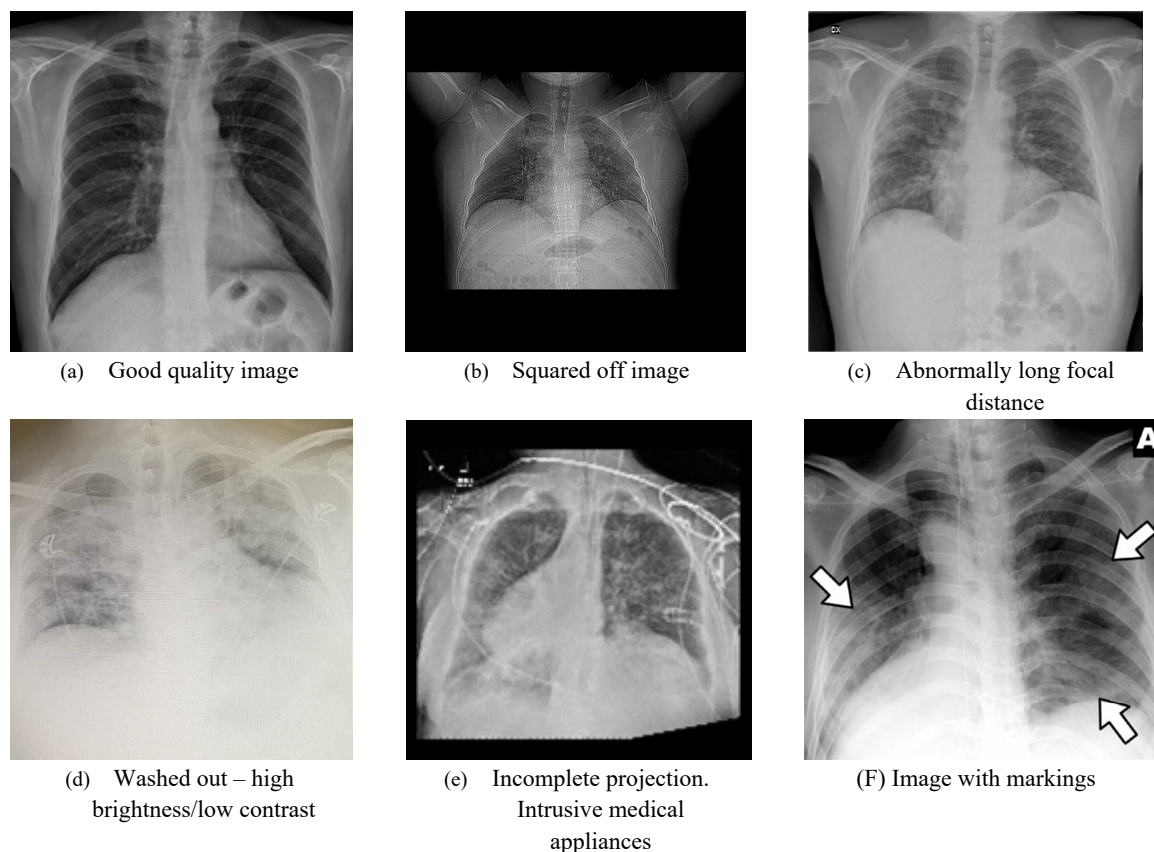
### III. DATASET DEVELOPMENT

#### A. Data sourcing

Large numbers of X-Ray images are available from a number of publicly accessible datasets. With the emergence of COVID-19 being very recent none of these large repositories contain any COVID-19 labelled data, thereby requiring that we rely upon at least two datasets for Normal, Pneumonia and COVID-19 source images.

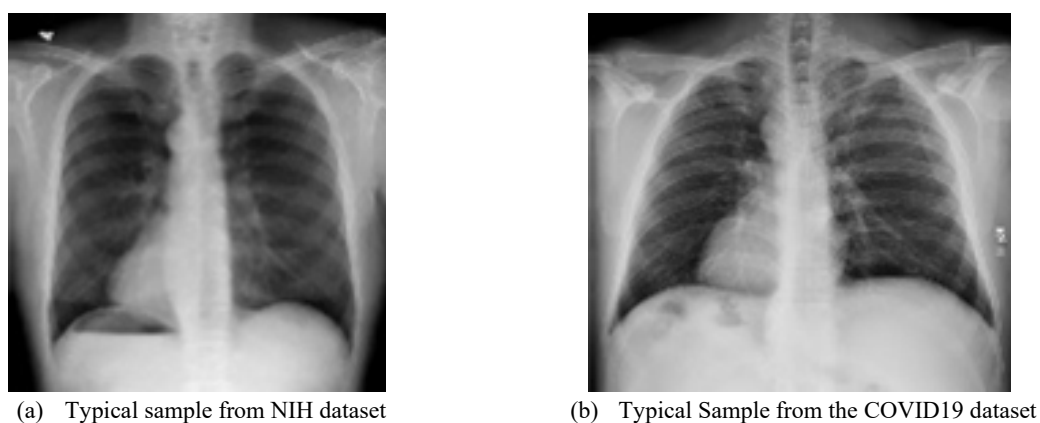
COVID-19 chest X-Rays were obtained from the publicly accessible COVID-19 Image Data Collection [11]. This collection has been sourced from websites and pdf format publications. Unsurprisingly, the images from this collection are of variable size and quality.

As of our most recent data that we have downloaded was on 12 April 2020. There was a total of 115 PA view images tagged as COVID-19 within this data set. The sizes of these images are highly variable with the smallest being 255 x 249 pixels and the largest being 4280 x 3520 pixels. Image contrast levels, brightness and subject positioning are all highly variable within this dataset. The data set mostly contains data for adult patients. Few examples of this variability are shown in Figure 3.



**Figure 3:** Different variations observed in the COVID-19 dataset.

National Institute of Health (NIH) Chest X-Ray [30] – a source of 112,120 anonymized labelled X-Ray images with 14 condition labels including Pneumonia and normal conditions. The images are of a similar size, [31] quality and aspect ratio to the typical images in the COVID-19 dataset with dimensions being uniformly 1024 x 1024 pixels in a portrait orientation. A comparison of images from the NIH dataset to the COVID-19 dataset is shown in Figure 4.



**Figure 4:** Visual comparison of samples from the NIH and COVID-19 datasets

Visually, the levels of intra dataset difference between images are similar. While the differences between the samples from two datasets are also not strongly apparent as it was evident between RSNA and COVID-19 datasets in Figure 2. We consequently chose this dataset as the source for pneumonia and normal condition X-Ray images.

Our findings in relation to available data sources is summarised in Table 1:

TABLE 1: SUMMARY OF DATA SOURCES CONSIDERED

Collection	Number of Images	Characteristics	Notes
COVID-19 Image Data Collection [11]	115 COVID-19 (PA)	Variable size, quality, contrast and brightness.	Only source of COVID-19 PA X-Ray images and used in this study.
Montgomery County X-Ray Set[32]	80 Normal 58 Tuberculosis	N/A	Not used due to too few normal images and no pneumonia labelled images
Kaggle RSNA Pneumonia Detection Challenge [28]	8851 Normal 9555 Lung Opacity 11821 No Lung Opacity/Not Normal	Very uniform. PA view differs to typical COVID-19.	Clear systematic visual differences to the COVID-19 collection, hence not used in this study
NIH Chest X-Ray[30]	322 Pneumonia 60361 No Finding	Uniformity similar to COVID-19 dataset. All images are 1024 x1024 in size.	Objectively similar in quality to the COVID-19 Image Data Collection. Used in this study.

### B. Data Sampling

In this study we aim to use real X-Ray data only and not considering creation and use of synthetic data at this stage. We also planned to use balanced dataset size for our model experiments. From the two source datasets COVID-19 and NIH, a master dataset containing COVID-19, Pneumonia and Normal images was created for our model development and testing purposes. The COVID-19 dataset was curated to remove images that were the wrong projection, poorly cropped, low resolution or where medical devices were a dominant feature (since these might draw the attention of the machine learning algorithms). This left us with 100 usable samples for COVID-19 dataset. Since the NIH dataset contains much more samples, we used down sampling technique to select matching numbers of samples for Normal and Pneumonia cases. The Pneumonia and Normal samples were selected from the downloaded NIH dataset randomly and excludes samples for juvenile patients. A small number of images selected in this way were found to contain medical devices (pacemakers) which we were concerned might draw the attention of the machine learning algorithms, so these images were discarded and replaced with a random selection until such devices were absent from the data set. The results of this procedure are summarised in Table 2.

TABLE 2: SAMPLED DATASET FOR EXPERIMENTS

Source	Condition	Number of source images	Number of curated images	Curation notes
COVID-19 Image Data Collection	COVID-19	115	100	Incorrectly labelled sample excluded Low quality images (low pixel count, poorly cropped etc.) excluded Samples with medical devices excluded
NIH Chest X-Ray	Pneumonia	322	100	100 samples chosen randomly Juvenile patients excluded Samples with medical devices
NIH Chest X-Ray	Normal	60361	200	200 samples chosen randomly Juvenile patients excluded Samples with medical devices

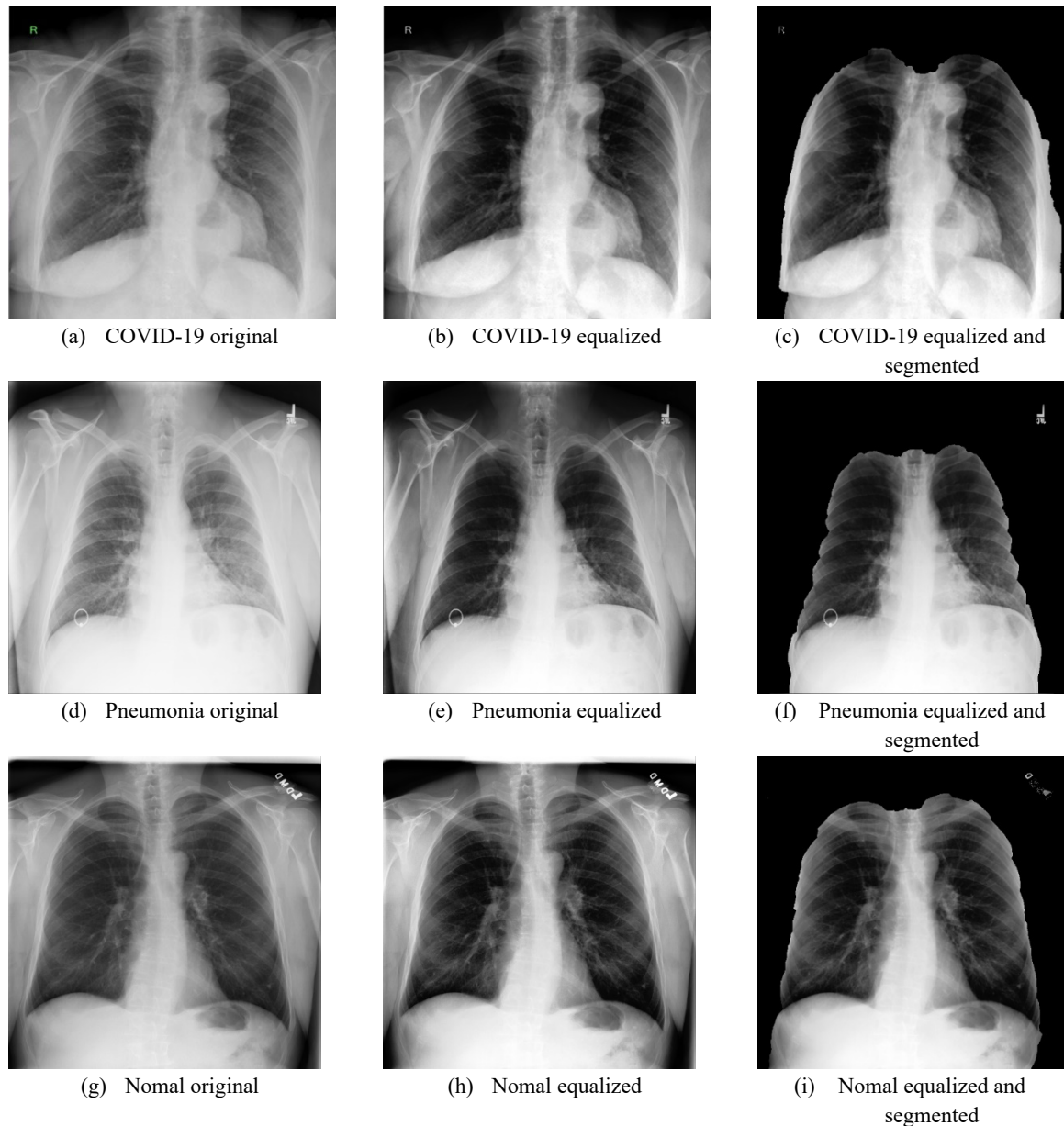
### C. Data Pre-Processing

#### 1) Equalization of Sampling Bias

We recognised that although we were using only two different datasets (COVID-19 and NIH) for our experiments; we were in fact relying upon data sourced from an unknown number and variety of X-Ray machines, exposure parameters and operator behaviour. Systematic image exposure and brightness differences for the datasets proved to be particularly concerning and several researchers have indicated that medical image analysis methods are highly sensitive to variations of image intensities [33] and, [34]. Research has shown that the feasibility of an automated analysis system requires that “the intensity of the same tissue should always be similar in one image and not vary with the location of the tissue”[34]. This principle when extrapolated to the many images utilised in machine learning algorithms, implies that all images in the sample data sets should have similar intensity for the same tissue *over the entire set of images* as well as within a single image.

Since the machine learning classifiers use a pixel array as a data source then any systematic difference in pixel intensity between the datasets would introduce sampling bias in the result. This would have the consequence of training the machine learning classifiers on systematic image histogram differences rather than the actual clinical image content of interest.

To minimize the effect of sampling bias, we applied histogram equalization to images using the OpenCV `equalizeHist` [35] function. As shown in Figure 5, the equalization process greatly improved the image brightness/contrast consistency across datasets. The same effect was observed within each data set. Subjectively, the authors can no longer easily tell which image has been drawn from which dataset purely on image brightness and contrast characteristics alone.



**Figure 5:** Results of equalization and segmentation on original samples for COVID-19, Pneumonia and Normal images.

## 2) Segmentation based noise reduction

Since the COVID-19 and Pneumonia affects the lung, we wanted to improve machine learning classification attention to the lung region of interest only. The aim is to remove the noise data such as collar bone, neckline, border, diaphragm etc. around the lung area. The natural way to achieve this was by cropping all but the lung area from the images sent to the classifier. The efficacy of this approach in medical image classification is supported in past studies [36]. A comprehensive review of lung area segmentation techniques may be found in



[37]. This study notes that intra-thoracic pathology, external objects such as jewellery and even subject position and posture all adversely affect the current lung segmentation algorithms and the problem of accurate lung segmentation from PA X-Rays is an open problem.

Our data sets show these problematic characteristics, and perfect automatic lung segmentation was therefore unlikely regardless of the technique applied. We applied the OpenCV GrabCut [38] function, which was previously used for CT scans [39] as a simple tool implement segmentation technique on our equalized images. We reasoned that the lung area could be considered the foreground of the X-Ray image. The segmented results for different image classes are shown in Figure 5.

### 3) Pre-processing Pipeline

Firstly, the pre-processing pipeline shown in Figure 6 is responsible for applying histogram equalization and the GrabCut algorithm to the sample images. Images are first read before being resized to 640 x 640 pixels. This initial resize is performed in order to optimize the performance of the downstream GrabCut process. We found by experimentation that GrabCut worked much better on the reversed image compared to the original image. Following the application of GrabCut, the image again reversed (back to normal) to complete the pre-process.

The detailed pre-processing pipeline is shown in figure 6.

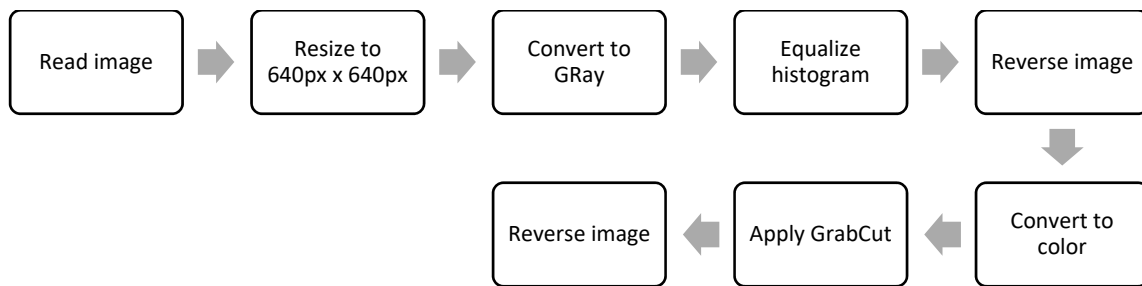


Figure 6: Data pre-processing pipeline.

## IV. MODEL DEVELOPMENT

### A. Classification Pipeline

The basic classification pipeline is shown in Figure 7, which reads in images sorted into subdirectories with subdirectory name used as the image classes. The images are resized into 224 x 224 pixels array before being randomly split into 80:20 training and testing data subsets. Image data was augmented only with a rotation operation with a range of 15.

Each tested model is then trained over 20+ epochs with precision, recall, training and testing accuracy and loss metrics are captured for further analysis.

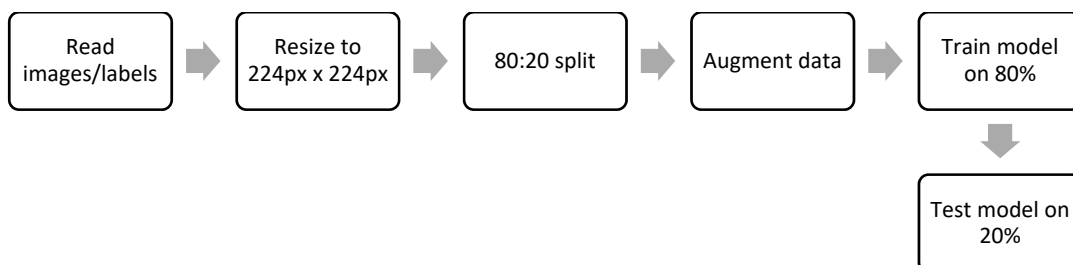


Figure 7: Classification pipeline.



### *B. Model Consideration*

One of the key objectives of this study was to achieve state-of-the-art classification results using publicly accessible data and “out-of-the-box” models with transfer learning to both compensate for the limited size of the sample data, and to accelerate the training process so that this could be reasonably performed on modest hardware.

For these reasons we have selected five deep learning models as classifiers for our experiments. Conveniently, these models are all available as part of the Keras API and each support transfer learning [44] in the form of supporting the pre-application to the model of the ImageNet [40] weights.

#### *1) VGG16 and VGG19*

VGG16 and VGG19 [15] are convolutional neural network (CNN) architectures with very small convolution filters (3x3) and a stride of 1 designed to achieve high accuracy in large-scale image recognition applications. The two implementations differ in depth of convolution/max-pooling and fully connected layers, with VGG16 having 16 layers in the base model and VGG19 having 19 layers.

#### *2) ResNet50*

The ResNet [16] CNN was developed as a means of avoiding the vanishing gradient problem inherent in deep neural networks by implementing a system of skip connections between layers – known as residual learning. This architecture results in a network that is more efficient to train, allowing for deeper networks to be designed that positively impact the model accuracy. ResNet50 is such a network with 50 layers implementing residual learning.

#### *3) Inception V3*

The Inception V3 [17] CNN aimed to improve utilization of computing resources inside the network by increasing the depth and width of the network whilst keeping computation operations constant. The designers of this network coined the term “inception modules” to describe an optimised network structure with skipped connections that is used as a building block. This inception module is repeated spatially by stacking with occasional max-pooling layers to reduce dimensionality to a manageable level for computation.

#### *4) Xception*

The Xception [18] CNN was developed by Google Inc. as an “extreme” version of the Inception model. The Inception modules described above are replaced with depth wise separable convolutions. This Xception was shown to outperform Inception on a large-scale image classification dataset (comprising 350 million images of 17,000 classes).

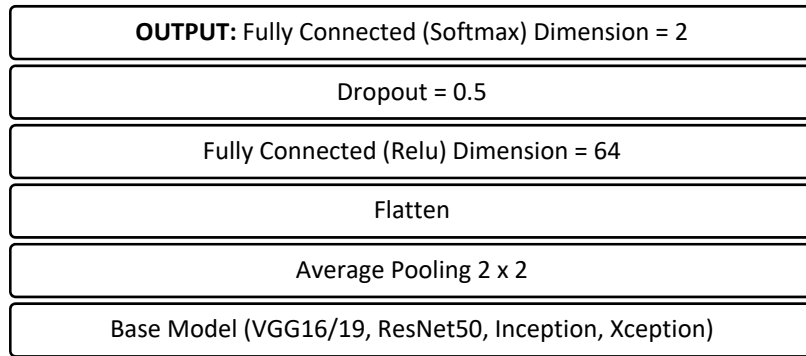
## V. MODEL ASSESSMENT

### *A. Computing Infrastructure*

All the experiments were performed using an off-the-shelf Lenovo P50 laptop with an Intel Core i7-6820HQ CPU @ 2.70GHz 4 Core CPU. The laptop hosts an NVIDIA Quadro M1000M GPU and an Intel HD Graphics 530 on-board graphics card. There was 32GB physical RAM installed.

### *B. Experiment Setup*

To ensure a consistent comparison across the models, each model was instantiated with a uniform output head as shown in Figure 8:



**Figure 8:** Head architecture of the proposed models.

Each experiment was performed 20 times with using 20 Epochs, Learning Rate =  $1e-3$  and a Batch Size of 8.

### C. Experiment Dataset

The master dataset was utilised for training and validation with each candidate classifier over 2 experiments per classifier as shown in Table 3.

Processed dataset is available for public use at <https://github.com/mhorry/SegmentedLungCXRs>

**TABLE 3:** DATASETS USED FOR EXPERIMENTS

Experiment ID	Experiment	Dataset
<b>Experiment 1</b>	Normal vs COVID-19 and Pneumonia.	200 x Normal vs 100 x COVID-19 100 x Pneumonia
<b>Experiment 2</b>	COVID-19 vs Pneumonia	100 x Pneumonia Vs 100 x COVID-19

### D. Comparative Results and Discussions

The experimental results are listed in Table 4. For each classifier we have noted average Precision (P), Recall (R) and F1 score over the 20 experimental runs.

Experiment 1 classified normal condition against a combined dataset of COVID-19 and Pneumonia. Experiment 2 classified COVID-19 vs Pneumonia.

The numbers provided in the Table 4 shows averages over 20 test cycles as percentages (%).

**TABLE 4:** PERFORMANCE RESULTS FOR TESTED MODELS

Experiment	VGG16			VGG19			ResNet50			InceptionV3			Xception		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
<b>Experiment 1: Normal vs. (COVID-19+ Pneumonia)</b>	82	80	80	83	80	80	70	67	66	68	65	61	69	57	48
<b>Experiment 2: COVID-19 vs, Pneumonia</b>	83	81	80	83	81	81	67	58	51	41	50	36	30	50	34

To understand these results it is useful to consider the learning curves for experiment 1 as shown in Figure 9.

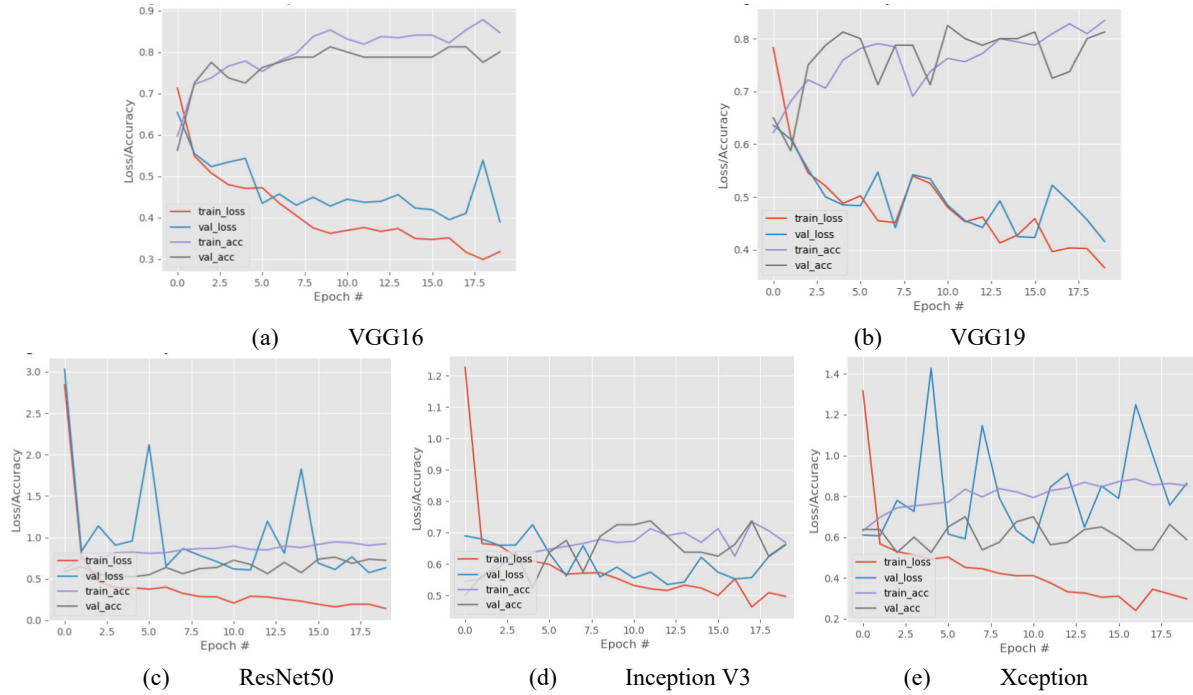


Figure 9: Learning curves for different models.

One interesting observation is that the deeper networks, being ResNet50, Inception and Xception exhibited poor performance compared to the shallower networks being VGG16 and VGG19. Noting that our sample sizes were small (200 samples for each class in Experiment 1 and 100 samples per class in Experiment 2 and the challenges inherent in training CNNs with small data sets as described by D’Souza et al. [41]. We suspect that we did not have enough X-Ray samples to train these deeper networks. Given the currently limited number of quality samples of COVID-19 X-Ray images, we believe that the VGG16 or VGG19 classifiers to be the most accurate off-the-shelf for classification of COVID-19 and Pneumonia from Normal condition, and for COVID-19 vs Pneumonia. Both VGG16 and VGG19 performed very well in Experiment 1 and both were successful in distinguishing between normal and COVID-19 or Pneumonia pathologies with an average F1 score of 80%. VGG19 slightly outperformed VGG16 in experiment 2 with a 1% improvement in F1 score for experiment 2 following an additional 20 training cycles indicated by the original VGG19 training curve for experiment 2.

As additional high quality COVID-19 X-Ray image sample become available we expect that the VGG-19 classifier to generally outperform the VGG-16 classifier. For this reason, we recommend the VGG19 classifier over the VGG16 classifier.

## VI. CONCLUSIONS

We have demonstrated that deep learning models for the automated identification of COVID-19 to be feasible using publicly available X-Ray image sample training/validation data and off-the-shelf deep learning algorithms.

We provided a pre-processing pipeline aimed to remove the sampling bias. Our results of around 80% for both recall and precision are reasonable firstly, given our small, and variable quality, COVID-19 X-Ray image sample size, and secondly that trained clinicians have an interpretative radiological error of 2 – 20% depending on the radiological investigation [42].

We found that both VGG16 and VGG19 classifiers provided good results within the experimental constraints of the small number of currently available COVID-19 X-Ray image samples. As the number of available COVID-19 X-Ray image samples increases we will be able to provide enough training data corpus to the deeper networks (ResNet50, Inception, and Xception) and thereby obtain better results from these classifiers. This will be the subject of our future work.

Despite our good results, we would urge great caution in the development of clinical diagnostic models using currently available COVID-19 X-Ray image data. The effect of a false positive diagnosis of COVID-19 on an individual is the isolation of the individual and their contract traces and the mental anguish and stress caused by both the prognosis and the social isolation. A false positive COVID-19 diagnosis could result in an inappropriate course of treatment. The effects of a false negative COVID-19 diagnosis would also be devastating for the individual if that diagnosis led to an inappropriate course of treatment, and the community since cautions against COVID-19 transmission may not be appropriately applied resulting in the further spread of the disease.

As a higher quality corpus of COVID-19 X-Ray image data becomes available, it may be possible to produce deep learning based clinical models for the fast diagnosis of COVID-19 as distinguished from similar conditions such as Pneumonia. Such a tool would prove invaluable in practice, where other diagnostic tests for COVID-19 are either unavailable or unreliable. As the COVID-19 spread progresses throughout remote and economically challenged locations an ability to diagnose COVID-19 from a readily available chest X-RAY image would help slow the spread of the disease and result in a better medical outcome for the population.

#### ACKNOWLEDGEMENT

We wish to thank IBM Australia for providing work time release for the 1<sup>st</sup> author to perform modelling and experimentation work and to contribute to the writing of this paper.

#### REFERENCES

- [1] WHO. "COVID-19 situation reports." <https://protect-au.mimecast.com/s/w3GmCJyBPjF1nJg2uQNMqk?domain=who.int> (accessed).
- [2] W. Wang *et al.*, "Detection of SARS-CoV-2 in Different Types of Clinical Specimens," *JAMA*, 2020, doi: 10.1001/jama.2020.3786.
- [3] C. Chen *et al.*, "SARS-CoV-2-Positive Sputum and Feces After Conversion of Pharyngeal Samples in Patients With COVID-19," *Annals of internal medicine*, 2020, doi: 10.7326/m20-0991.
- [4] T. Ai *et al.*, "Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," *Radiology*, vol. 0, no. 0, p. 200642, 2020, doi: 10.1148/radiol.2020200642.
- [5] 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," (in en), <https://doi.org/10.2214/AJR.20.22969>, review-article 2020-02-28 2020, doi: 10.2214/AJR.20.22969.
- [6] J. Chen *et al.*, "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study," *medRxiv*, p. 2020.02.25.20021568, 2020, doi: 10.1101/2020.02.25.20021568.
- [7] S. Wang *et al.*, "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)," *medRxiv*, p. 2020.02.14.20023028, 2020, doi: 10.1101/2020.02.14.20023028.
- [8] E. Livingston *et al.*, "Sourcing Personal Protective Equipment During the COVID-19 Pandemic," *JAMA*, 2020, doi: 10.1001/jama.2020.5317.
- [9] L. J. M. Kroft, L. van der Velden, I. H. Giron, J. J. H. Roelofs, A. de Roos, and J. Geleijns, "Added Value of Ultra-low-dose Computed Tomography, Dose Equivalent to Chest X-Ray Radiography, for Diagnosing Chest Pathology," (in eng), *J Thorac Imaging*, vol. 34, no. 3, pp. 179-186, May 2019, doi: 10.1097/rti.0000000000000404.
- [10] H. Y. F. Wong *et al.*, "Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients," (in EN), <https://doi.org/10.1148/radiol.2020201160>, research-article 2020-03-27 2020, doi: <https://doi.org/10.1148/radiol.2020201160>.
- [11] J. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," *arXiv.org*, 2020.
- [12] "ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection." <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection> (accessed).
- [13] R. Yamashita, M. Nishio, R. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4, pp. 611-629, 2018, doi: 10.1007/s13244-018-0639-9.
- [14] M. Mazurowski, M. Buda, A. Saha, and M. Bashir, "Deep learning in radiology: an overview of the concepts and a survey of the state of the art," *arXiv.org*, 2018.
- [15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv.org*, 2015.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *arXiv.org*, 2015.
- [17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," vol. 2016-, ed: IEEE, 2016, pp. 2818-2826.
- [18] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *arXiv.org*, 2017.
- [19] P. Pisani *et al.*, "Screening and early diagnosis of osteoporosis through X-ray and ultrasound based techniques," *World journal of radiology*, vol. 5, no. 11, pp. 398-410, 2013, doi: 10.4329/wjr.v5.i11.398.
- [20] M. A. Al-Antari, M. A. Al-Masni, M.-T. Choi, S.-M. Han, and T.-S. Kim, "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification," *International Journal of Medical Informatics*, vol. 117, pp. 44-54, 2018, doi: 10.1016/j.ijmedinf.2018.06.003.

- [21] M. A. Speidel, B. P. Wilfley, J. M. Star-Lack, J. A. Heanue, and M. S. Van Lysel, "Scanning-beam digital x-ray (SBDX) technology for interventional and diagnostic cardiac angiography," *Medical Physics*, vol. 33, no. 8, pp. 2714-2727, 2006, doi: 10.1118/1.2208736.
- [22] N. Kanwal, A. Girdhar, and S. Gupta, "Region Based Adaptive Contrast Enhancement of Medical X-Ray Images," ed: IEEE, 2011, pp. 1-5.
- [23] M. R. Arbabshirani, A. H. Dallal, C. Agarwal, A. Patel, and G. Moore, "Accurate segmentation of lung fields on chest radiographs using deep convolutional networks," in *Medical Imaging 2017: Image Processing*, E. D. Angelini and M. A. Styner, Eds., 2017, vol. 10133: SPIE, doi: 10.1117/12.2254526. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85020298165&doi=10.1117%2f12.2254526&partnerID=40&md5=dde9a12a65da38c5a3ca325721d77192>
- [24] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection," *arXiv.org*, 2020.
- [25] E. El-Din Hemdan and M. Shouman, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," *arXiv.org*, 2020.
- [26] M. Chowdhury *et al.*, "Can AI help in screening Viral and COVID-19 pneumonia?," *arXiv.org*, 2020.
- [27] L. Wang and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images," *arXiv.org*, 2020.
- [28] "RSNA Announces Pneumonia Detection Machine Learning Challenge," ed: PR Newswire Association LLC, 2018.
- [29] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," (in En), *Journal of Big Data*, OriginalPaper vol. 6, no. 1, pp. 1-54, 2019-03-19 2019, doi: doi:10.1186/s40537-019-0192-5.
- [30] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," vol. 2017-, pp. 3462-3471, 2017, doi: 10.1109/CVPR.2017.369.
- [31] "National Institutes of Health - NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community," ed: Normans Media Ltd., 2017.
- [32] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wang, P.-X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quantitative imaging in medicine and surgery*, vol. 4, no. 6, pp. 475-477, 2014, doi: 10.3978/j.issn.2223-4292.2014.11.20.
- [33] U. Vovk, F. Pernus, and B. Likar, "A Review of Methods for Correction of Intensity Inhomogeneity in MRI," *IEEE Transactions on Medical Imaging*, vol. 26, no. 3, pp. 405-421, 2007, doi: 10.1109/tmi.2006.891486.
- [34] S. Song, Y. Zheng, and Y. He, "A review of Methods for Bias Correction in Medical Images," *Biomedical Engineering Review*, vol. 1, no. 1, 2017-09-15 2017, doi: 10.18103/bme.v3i1.1550.
- [35] "OpenCV: Histograms - 2: Histogram Equalization." [https://docs.opencv.org/master/d5/daf/tutorial\\_py\\_histogram\\_equalization.html](https://docs.opencv.org/master/d5/daf/tutorial_py_histogram_equalization.html) (accessed).
- [36] A. V. Rabinovich, Andrea & Belongie, Serge., "Does Image Segmentation Improve Object Categorization? | Request PDF," 2020, doi: <http://dx.doi.org/>.
- [37] S. Candemir and S. Antani, "A review on lung boundary detection in chest X-rays," *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 4, pp. 563-576, 2019, doi: 10.1007/s11548-019-01917-1.
- [38] "OpenCV: Interactive Foreground Extraction using GrabCut Algorithm." node. [https://docs.opencv.org/3.4/d8/d83/tutorial\\_py\\_grabcut.html](https://docs.opencv.org/3.4/d8/d83/tutorial_py_grabcut.html) (accessed).
- [39] S. Zhang, Y. Zhao, and P. Bai, "Object Localization improved GrabCut for Lung Parenchyma Segmentation," *Procedia Computer Science*, vol. 131, pp. 1311-1317, 2018, doi: 10.1016/j.procs.2018.04.330.
- [40] J. Deng, W. Dong, R. Socher, L. Li, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 20-25 June 2009 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.
- [41] R. N. D'Souza, P.-Y. Huang, and F.-C. Yeh, "Structural Analysis and Optimization of Convolutional Neural Networks with a Small Sample Size," *Scientific reports*, vol. 10, no. 1, pp. 834-834, 2020, doi: 10.1038/s41598-020-57866-2.
- [42] A. P. Brady, "Error and discrepancy in radiology: inevitable or avoidable?," (in eng), *Insights Imaging*, vol. 8, no. 1, pp. 171-182, 2017, doi: 10.1007/s13244-016-0534-1.
- [43] A Ulhaq, A Khan, D Gomes, M Paul, "Computer vision for COVID-19 control: a survey," *engrXiv*, 1-24, 2020
- [44] Shermin T., Teng S.W., Murshed M., Lu G., Soheli F., Paul M. (2019) Enhanced Transfer Learning with ImageNet Trained Classification Layer. In: Lee C., Su Z., Sugimoto A. (eds) *Image and Video Technology. PSIVT 2019. Lecture Notes in Computer Science*, vol 11854. Springer, Cham