Efficient-CovidNet : Deep Learning Based COVID-19 Detection From Chest X-Ray Images

Yash Chaudhary¹, Manan Mehta¹, Raghav Sharma², Deepak Gupta^{1,3}, Ashish Khanna^{1,3}, and Joel J. P. C. Rodrigues^{3,4}

¹Maharaja Agrasen Institute of Technology, Delhi, India

²University School of Information, Communication and Technology, Delhi, India

³Federal University of Piauí, Teresina – PI, Brazil

⁴Instituto de Telecomunicações, Portugal

yash612@gmail.com, manan.mehta2@gmail.com, sharma.raghav103@gmail.com, deepakgupta@mait.ac.in, ashishkhanna@mait.ac.in, joeljr@ieee.org

Abstract— The COVID-19 pandemic has wreaked havoc all over the world. The rising number of cases have overburdened healthcare systems even in the most developed countries. To ease the burden on healthcare systems a quick and efficient testing technique is needed. Currently, the RT-PCR testing is done with time consuming and laborious an alternative is a detection from Chest X-Ray images. It has been discovered in published studies that Chest X-Rays of COVID-19 patients have specific malformations that can be used to identify a positive case. Inspired by the work done on "COVID-Net" by Linda Wang, Zhong Qiu Lin and Alexander Wong, a Deep Learning approach to detect coronavirus from Chest X-Ray images is used in this study. To surpass previous results the EfficientNet Convolutional Neural Network (CNN) model is proposed. This model not only achieves +2% accuracy, but it also attains higher sensitivity and Positive Predictive Values. The study uses the open source COVIDx dataset. It has approximately 14,000 X-Ray images. To the best of authors' knowledge, this dataset contains the largest number of COVID-19 positive cases. The study offers a Deep Learning approach contributing to create an efficient COVID-19 detector that can be used in the real world.

Keywords — COVID-19; Convolutional Neural Network; Deep Learning; X-Ray images; SARS-CoV-2.

I. INTRODUCTION

The COVID-19 pandemic has brought the world to a standstill. From the crumbling economy to rising death tolls, this pandemic is affecting everyone. It's important to contain the spread to minimize the risk of infection. This can be done through widespread testing and contact tracing. Countries like Singapore, Taiwan, South Korea have shown proof of the efficacy of widespread testing and tracing, these countries have been successful in flattening their curves early on. Thus it becomes necessary to intensify the screening and tracing techniques. In this study we focus on how the COVID-19 screening can be expedited. The chief method used for detecting Coronavirus positive cases is RT-PCR testing, which can detect SARS-CoV-2 RNA from respiratory samples. While the abovementioned test for detecting COVID-19 cases is the most

accurate and most reliable due to its its high specificity, it is a very time-consuming, burdensome and a tricky manual process. Moreover, RT-PCR tests have been reported to have variable sensitivity and fluctuating positive rate. Some recent findings reveal that these tests show variable positive rate depending on how the respiratory samples were accrued. Some studies also indicate a decreased positive rate with time after symptom onset [3, 4]. Hence we promote an alternative screening technique in which chest X-Ray images can be efficiently used to predict COVID-19 viral infection. Early studies have claimed that abnormalities in the chest radiography can be correlated to COVID-19 and because of this it can be our most prominent tool for COVID-19 screening [5, 6]. The various benefits of implementing CXR (Chest X-Ray) imaging for COVID-19 screening, especially in resource-constrained regions and heavily-affected zones.

- 1) Rapid Testing
- 2) Approachable
- 3) Availability
- 4) Portability

Being faster, more convenient, and more cost-efficient than the standard test used for the screening of COVID-19, CXR (Chest X-Ray) imaging can become one of the most beneficent methods of testing and expedite the other processes. Moreover, a few have indicated that as the COVID-19 pandemic advances, that there will be increased dependence on portable X-Ray machines due to the benefits mentioned earlier. In this study we leverage a popular Deep Learning network (Convolutional Neural Network) to classify a given Chest X-Ray image according to the three classes :- COVID-19, pneumonia, normal case (healthy patient). Figure 1 and Figure 2 show examples of pneumonia and COIVD-19 infections respectively.

Convolutional Neural Network (CNN), in Computer Vision, have been used for image classification tasks. Previously, numerous researches of medicinal imagery have utilized CNN to register a more reliable performance. This study includes the following main contributions:

© IEEE 2021. This article is free to access and download, along with rights for full text and data mining, re-use and analysis.

- Application of state of the art CNN EfficientNet on Chest X-Ray Images for classification. It was observed that EfficientNet shows some of the best results.
- A comparative study between the most promising solutions available in the literature for Chest X-Ray Image classification.



Fig. 1. Illustration of Non-COVID19 infection.



Fig. 2. COVID-19 infection in COVIDx dataset.

The rest of the paper is organised as follows. Section II elaborates on the related work on deep learning for COVID-19 detection. A presentation of the dataset used for COVID-19 detection is presented in Section III. Section IV describes how deep learning is used for image classification and the model (EfficientNet) used for this task is explained. A discussion for choosing EfficientNet and its architecture is considered. Section V and VI details the solution deployment and evaluation metrics used in the study. Section 7 present the obtained results and analysis showing the efficacy of EfficientNet model. Detailed performance comparisons with other models trained on the same (COVIDx) Dataset are considered. Finally, the paper is concluded in Section VIII considering suggestions for further works in the topic.

II. RELATED WORK

The literature demonstrates that the abnormalities in chest radiography can be associated with COVID-19 [5,6]. It has opened the door for researchers to detect COVID-19 from Chest X-Ray images. Due to the convenience and affordability of this novel method many researchers have conducted studies for effective screening of COVID-19 from X-Ray images. The evaluation of different families of CNNs [15] and two step modular procedure [16] has been used on COVIDx dataset. A novel CNN architecture specifically designed for the task of COVID-19 detection have been implemented in literature [1]. Aforementioned literature focuses on the exploration of deep convolutional neural networks, due to the meaningful successes in the field of image classification.

This work has been inspired by a new architecture of CNN, called "COVID-Net" is discovered [1]. In this work the neural architecture search algorithm leveraged to discover the CNN architecture. The model classifies CXR images into normal, pneumonia, and COVID-19. Distinctively from the previous work, It uses a much larger dataset consisting of 13,975 CXR images across 13,870 patient cases [1]. The method results with an accuracy of 93.3 percent overall and sensitivity of 96.0 percent for COVID-19. The proposed work was motivated by further improving the accuracy and efficiency of "COVID-Net".

III. COVIDX DATASET



Fig. 3. Example of Covidx Dataset.

The dataset chosen to train and test our model is the COVIDxv3 (version 3) [1]. It is by far the most comprehensive and largest dataset for Chest X-Ray images. It contains a total of 13,975 Chest X-Ray Images images across 13,870 patients. The dataset has 3 categories: normal, pneumonia (Non-COVID-19) and COVID-19. Making it a multi classification problem.

The dataset is completely open source, it is freely available on github. It is made up of 5 different public data resources: COVID-19 Image Data Collection [7], COVID-19 Chest X-Ray Dataset Initiative [8], ActualMed COVID-19 Chest X-Ray Dataset Initiative [9], RSNA Pneumonia Detection Challenge dataset, COVID-19 radiography database.

IV. RESEARCH METHODOLOGY

This section explains why deep learning has been successful for image classification tasks. Further it describes in detail the model used for COVID-19 classification, including the reason for choosing the model, its architecture and its scaling methodology.

A. Method

Deep Learning has proved to be very effective in image classification due to the recent advances in Deep Learning architectures known as, Convolutional Neural Networks (CNN). CNN's became famous in 2012 when AlexNet [19] won "ImageNet Large Scale Visual Recognition challenge" [18]. Since then the Deep Learning community has continually improved CNN architectures, some of the famous CNN architectures are VGG, GoogleNet, Resnet and EfficientNet. CNN's are good at image classification tasks because of their ability to detect location invariant features. When we train a CNN it automatically learns filters that act as feature extractors. These learned filters help the model to understand and distinguish between different classes of images input to the model. Figure 4 depicts how a CNN extracts features from an image and uses it for classification.



Fig. 4. Convolutional Neural Network for Image Classification.

This study leverages CNN architecture to classify CXR (Chest X-Ray) images into three classes Normal, Pneumonia (Non COVID-19) and COVID-19. We use EfficientNet CNN architecture [10] for this purpose.

B. Efficient Net

EfficientNets [10] are a family of models which are one of the most efficient yet most accurate ConvNets till date. In the study [10] authors proposed a new ConvNet scaling technique which uniformly scales width, resolution and depth of ConvNets to achieve better accuracy and efficiency. Furthermore, they used this new scaling technique to obtain a new set of models called EfficientNets.

There are 8 original EfficientNet models EfficientNet-B0 - EfficientNet-B7. The EfficientNet-B0 is the baseline model and all the other models (B1 - B7) are scaled versions of it (i.e as we go from B0 to B7 the number of parameters and architectural complexity of model increases). This research

focuses only on models EfficientNet-B0 - EfficientNet-B3, since model's architectural complexity and size was imperative to us. After experimentation it was found that EfficientNet-B1 performed best, giving best accuracy and efficiency.



Fig. 5. Model Size vs Imagenet Accuracy (Adapted from [10]).

The intuition behind using EfficientNet was that these models have relatively fewer parameters, low FLOPS cost, they have state of the art results in image classification and they transfer well to other datasets.

To quote the authors [10] : "In particular, our EfficientNet-B7 achieves state-of-the-art 84.4 % top-1 / 97.1 % top-5 accuracy on ImageNet, while being 8.4 times smaller and 6.1 times faster on inference than the best existing ConvNet. Our EfficientNets also transfer well and achieve state-of-the-art accuracy on CIFAR-100 (91.7 %), Flowers (98.8 %), and 3 other transfer learning datasets, with an order of magnitude fewer parameters."

The hope was that since EfficientNets had been so successful on other datasets, this success could replicated on COVIDx dataset as well. And indeed EfficientNet proved to be great a choice, not only we got better results than COVID-net [1]. Our model only has 6.57 million parameters which is 44.08 % less than COVID-net.

C. Network Architecture

The baseline architecture, EfficientNet-B0 was developed by Neural Architecture Search algorithm that optimized for accuracy and FLOPS. The authors used a similar search space to [11] and hence the obtained architecture is similar to Mnas-Net [11]. Given below is the EfficientNet-B0 architecture. The MBConv block is Inverted Residual Block used in MobileNetV2 [12] with a Squeeze and excitation optimization.

D. Compound Scaling

ConvNets can be scaled across 3 dimensions :- width, depth and resolution. Previous efforts to scale ConvNets had focused only on scaling one of the three dimensions. [10] proves that balancing all the three dimensions of the network - depth,

stage	Operator	Resolution	Channels	Layers
1	Conv3x3	224x224	32	1
2	MBConv1,3x3	112x112	16	1
3	MBConv6,3x3	112x112	24	2
4	MBConv6,5x5	56x56	40	2
5	MBconv6,3x3	28x28	80	3
6	MBConv6,5x5	14x14	112	3
7	MBConv6,5x5	14x14	192	4
8	MBConv6,3x3	7x7	320	1
9	Conv1x1 with pooling & FC	7x7	1280	1

TABLE I EfficientNet-B0 Architecture.

width and resolution, while scaling gives more efficient and accurate models than ever before. Figure 7 shows different types of scalings.



Fig. 6. Different types of Scaling.

The intuition behind compound scaling is : "Intuitively, for higher resolution images, we should increase network depth, such that the larger receptive fields can help capture similar features that include more pixels in bigger images. Correspondingly, we should also increase network width when resolution is higher, in order to capture more fine-grained patterns with more pixels in high resolution images. These intuitions suggest that we need to coordinate and balance different scaling dimensions rather than conventional singledimension scaling." [10]

$$depth: d = \alpha^{\phi} \tag{1}$$

width :
$$w = \beta^{\phi}$$
 (2)

$$resolution: r = \gamma^{\phi} \tag{3}$$

s.t
$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2, \alpha \ge 1, \beta \ge 1, \gamma \ge 1$$
 (4)

In the equations above, ϕ is the compound coefficient, which the user specifies according to how much resources are at hand for scaling.

 α , β and γ are constants that indicate how these resources will be split among the depth, width, resolution of the ConvNet. We can obtain α , β and γ through grid search and after that user specified compound coefficient ϕ can be used to scale a baseline model.

V. TRAINING AND DEPLOYMENT DETAILS

EfficientNet-B1 pretrained on Imagenet Dataset [17] was used. Since the EfficientNet-B1 was previously trained on another classification problem (Imagenet has 1000 classes), the last Fully Connected layer which is also known as top layer was removed and replaced with a Dropout layer followed by a Fully connected layer with 3 nodes (i.e. for three classes normal, pneumonia, COVID-19). We use Dropout and Gaussian-Dropout layer as it is a powerful method for regularisation. It works by randomly dropping neurons while training. The activations of the" dropped-neurons" are ignored in the forward pass and these neurons have no contributions to weight update in the backward pass where the Dropout layer gave better results.

The hyper-parameters for training are the input image of size 512 x 512, Augmentation: None, Optimizer: Adam with AMSGRAD, Learning Rate: 0.0001, Loss: categorical Cross Entropy, No. of Classes: 3 (Normal, Pneumonia, COVID-19), Batch Size: 16, Epoch: 12. With 12 epochs, Training time was 2 hours (8-10 min per epoch) on Kaggle Kernel with 4 core CPU,16 GB RAM and Nvidia Tesla P100 GPU configuration.

VI. EVALUATION METHODS

When it comes to classification tasks in Machine learning, we use some specific metrics like precision, recall and F1 score because accuracy is not enough. Machine learning in the medical domain has a different perspective and to convince the medical field that AI can be used, for that we need a proper evaluation method to validate our results.

Sensitivity : Sensitivity measures the proportion of positive real cases which have been predicted as positive. Sensitivity is often termed as Recall, the sensitivity score indicates, what percentage of actual positive cases were correctly predicted by the model. For example, the test set had 100 COVID-19 images, so the sensitivity score for COVID-19 class indicates what percentage of 100 COVID-19 patients were correctly predicted by our model (EfficientNet-B1). Mathematically, it is a ratio between True positives and sum of True positives + False Negatives. The equation (5) describes sensitivity.

$$Sensitivity = \frac{Truepositives(TP)}{Truepositives(TP) + FalseNegatives(FN)}$$
(5)

Positive prediction value : It is defined as a ratio of true positives, and sum of True Positives + False Positives, equation (6). It is a measurement of how relevant positive result, It tells us how many predictions out of all the predictions were correct.

$$PPV = \frac{Truepositives(TP)}{Truepositives(TP) + FalsePositives(FP)}$$
(6)

VII. EXPERIMENTS AND RESULTS ANALYSIS

This section explains the various experiments conducted while training the CNN model, we throw light on various hyper-parameter optimizations. We discuss the challenges we faced while training and their solutions

A. Experiments

Various models and hyper-parameters were tested for obtaining best results. We fine-tuned EfficientNet-B0 to B3 as we wanted model to have small size. EfficientNet-B1 showed a perfect balance between testing accuracy and PPV/Sensitivity, leading us to make it our main model for further testing. We experimented by modifying the EfficientNet-B1 architecture by adding extra layers to the model but it was observed that more layers are giving uncertain PPV/Sensitivity. So we decided to stick with, layer of dropout followed by a Dense layer for image classification.

For training we tried Classical Data augmentation techniques like horizontal flip, vertical flip, rotation, zoom and progressive resizing but these techniques were tampering our results. So we decided to just normalize and resize images. We trained the model with various input sizes like 224 x 224, 416 x 416, 512 x 512. We got best results for input size of 512 x 512. We experimented leading us to chose a learning rate (LR) of 0.0001 with decay using ADAM optimizer with AMSGRAD. We used categorical cross-entropy loss function.

Another problem we faced while training was the imbalance in dataset. In the training dataset there were 7966 images of normal cases, 5459 images of pneumonia (Non-COVID 19) cases and only 473 images of COVID-19 cases. To stop our model from being biased against COVID-19 class, we thought that using focal-loss loss function would be great but it wasn't up to the mark, so for dealing with imbalanced dataset we used class weights. Class weights are a way of specifying to the model what percentage of loss should be propagated back for each class. This in turn effects the weight update for each class. For the majority class we can assign small class weights and for the minority class we can assign large class weights, we do this to limit the magnitude of weight update in favour of majority class. We applied high class weights to the COVID-19 class whereas low weights to the other two classes. The intuition was that there are very few examples of COVID-19 class so we needed to penalize weight updates of other two classes.

TABLE II Comparison of Model size.

Model	Parameters (Millions)
VGG-19	20.37
Resnet-50	24.97
COVIDNet	11.75
EfficientNet-B1	6.6

B. Results

The evaluation of the proposed deep neural network is done as the test accuracy, positive predictive value (PPV) with sensitivity for 2 infection types on the Industry benchmark COVID-x dataset. The metrics of the proposed model displayed in Table 3 and 4. The final-test accuracy is at 95 per cent whereas CovidNet was 93.3 per cent, thus showing the effective design and architecture of the proposed model meeting along with task, data, and operational requirements of disease detection. The deep neural network model can be applied to continuously increasing, and updated data as new cases are added. The sensitivity at 100 percent for the COVID-19 cases is judged good since no case of COVID-19 is missed. The PPV is also high at 94.3 per cent, showing very few false-positive case detections. The low false positive detections lower the need for wasteful PCR testing and patient care, thus saving the already stressed healthcare system.

TABLE III Test Accuracy.

Test Accuracy(%)				
Architecture	Test Accuracy			
EfficientNet-B1	95			
VGG-19	83.0			
Resnet-50	90.6			
COVIDNet	93.3			

TABLE IV Sensitivity Comparison.

Sensitivity(%)							
Architecture	Normal	Non-COVID-19	COVID-19				
EfficientNet-B1	95.5	94.6	100				
VGG-19	98.0	90.0	58.7				
Resnet-50	97.0	92.0	83.0				
COVIDNet-CXR4-A	94.0	94.0	95.0				
COVIDNet-CXR4-B	96.0	92.0	93.0				
COVIDNet-CXR4-C	95.0	89.0	96.0				

 TABLE V

 POSITIVE PREDICTIVE VALUE (PPV) COMPARISON.

PPV (%)							
Architecture.	Normal	Non-COVID-19	COVID-19				
EfficientNet-B1	93.5	96.9	94.3				
VGG-19	83.1	75.0	98.4				
Resnet-50	88.2	86.8	98.8				
COVIDNet-CXR4- A	91.3	93.1	99.0				
COVIDNet-CXR4- B	88.9	93.9	98.9				
COVIDNet-CXR4- C	90.5	93.7	96.0				

For comparison, the test results are bench-marked to the ongoing open-source COVIDNet project, model COVIDNet-A on COVIDx (100 samples of COVID-19) [2], and results are encouraging. The accuracy and the sensitivity result of the proposed model are better than COVIDNet model project

results. In contrast, the positive predictive value results for the proposed model are found to be better than COVIDNet models, in case of Normal and Pneumonia, except the case of COVID-19 where the EfficientNet-B1 model fall short by just 5.7 percentage from the COVIDNet-CXR4-A model. The COVIDNet model architecture is documented better than models VGG-19 [13] and ResNet-50 [14] as illustrated in [1]; thus the partial comparative advantage of the proposed deep neural network in his study is quite encouraging.



Fig. 7. Confusion Matrix represented in a Heat map.

VIII. CONCLUSION AND FUTURE WORKS

In Conclusion, we used EfficientNet-B1 for COVID-19 cases detection from Chest X-Ray images. We used the opensource COVIDx3 dataset. Our study proves the efficacy of the EfficientNet model on this task. Our results indicate that EfficientNet-B1 has better sensitivity and PPV than many models. It shows better results than the original COVIDNet models, which were specifically generated for this task through a Neural architecture search algorithm. We are proud of the fact that our model has sensitivity of 100% for the COVID-19 class. This means that our model correctly predicted all the COVID-19 patients in the test set. To the best of our knowledge, EfficientNet-B1 shows best results.

Future works may include continually improving PPV and sensitivity by training on newly collected data and experimenting the software in real world scenarios. Experimenting the model in real world exposes the model to data that it has never scene before. If the model is robust enough to keep the same level of accuracy as exhibited in the lab, then this would help in clearing the air of skepticism among healthcare professional to such emerging technologies.

ACKNOWLEDGEMENTS

This work is partially funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the Project UIDB/50008/2020; and by Brazilian National Council for Scientific and Technological Development - CNPq, via Grant No. 309335/2017-5.

REFERENCES

- L. Wang, Z. Q. Lin and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," arXiv preprint arXiv:2003.09871, 2020.
- [2] L. Wang, Lindawangg/COVID-Net. [online] GitHub. Available at: https://github.com/lindawangg/COVID-Net;, Accessed 1 July 2020.
- [3] Y. Yang et al, "Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019ncov infections," medRxiv:2020.02.11.20021493.
- [4] P. Wikramaratna, R. S. Palton, M. Ghafari, J. Lourenco "Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR," medRxiv; 2020, DOI: 10.1101/2020.04.05.20053355.
- [5] M.-Y. Ng et al, "Imaging profile of the COVID-19 infection: Radiologic findings and literature review,". Radiol. Cardiothorac. Imaging 2, 2020.
- [6] C. Huang et al, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, china," The Lancet 395, 497–506, 2020.
- [7] J. P. Cohen, P. Morrison and L. Dao, "COVID-19 image data collection," arXiv 2003.11597, 2020.
- [8] A. Chung, Figure 1 "COVID-19 chest x-ray data initiative," https://github.com/agchung/Figure1-COVID-chestxray-dataset, 2020.
- [9] A. Chung, "Actualmed COVID-19 chest x-ray data initiative," https://github.com/agchung/Actualmed-COVID-chestxray-dataset, 2020.
- [10] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," arXiv preprint arXiv:1905.11946, 2019.
- [11] M. Tan et al, "MnasNet: Platform-aware neural architecture search for mobile," CVPR, 2019.
- [12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: "Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474.
- [13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556,2014.
- [14] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [15] E. Luz et al, "Towards an effective and efficient deep learning model for covid-19 patterns detection in x-ray images," arXiv:2004.05717, 2020.
- [16] S. Khobahi, C. Agarwal and M. Soltanalian, "Coronet: A deep network architecture for semi-supervised task-based identification of covid-19 from chest x-ray images," medRxiv DOI: 10.1101/2020.04.14.20065722, 2020.
- [17] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," IEEE Computer Vision and Pattern Recognition (CVPR), 2009.
- [18] O. Russakovsky et al, "ImageNet Large Scale Visual Recognition Challenge," IJCV, 2015.
- [19] A. Krizhevsky, I. Sutskever and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Neural Information Processing Systems. 25. 10.1145/3065386, 2012.
- [20] C. Chakraborty, A. Banerjee, L. Garg, J. J. P. C. Rodrigues, "Internet of Medical Things for Smart Healthcare: Covid-19 Pandemic", Springer, Singapore, Series Studies in Big Data, Vol. 80, 2021, ISSN: 978-981-15-8096-3, eBook ISBN: 978-981-15-8097-0, DOI: 10.1007/978-981-15-8097-0.