



Deep efficient-nets with transfer learning assisted detection of COVID-19 using chest X-ray radiology imaging

Hiba Mzoughi¹ · Ines Njeh^{1,2} · Mohamed Ben Slima^{1,3} · Ahmed BenHamida¹

Received: 14 February 2022 / Revised: 28 February 2023 / Accepted: 10 March 2023 /

Published online: 27 March 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Corona Virus (COVID-19) could be considered as one of the most devastating pandemics of the twenty-first century. The effective and the rapid screening of infected patients could reduce the mortality and even the contagion rate. Chest X-ray radiology could be designed as one of the effective screening techniques for COVID-19 exploration. In this paper, we propose an advanced approach based on deep learning architecture to automatic and effective screening techniques dedicated to the COVID-19 exploration through chest X-ray (CXR) imaging. Despite the success of state-of-the-art deep learning-based models for COVID-19 detection, they might suffer from several problems such as the huge memory and the computational requirement, the overfitting effect, and the high variance. To alleviate these issues, we investigate the Transfer Learning to the Efficient-Nets models. Next, we fine-tuned the whole network to select the optimal hyperparameters. Furthermore, in the preprocessing step, we consider an intensity-normalization method succeeded by some data augmentation techniques to solve the imbalanced dataset classes' issues. The proposed approach has presented a good performance in detecting patients attained by COVID-19 achieving an accuracy rate of 99.0% and 98% respectively using training and testing datasets. A comparative study over a publicly available dataset with the recently published deep-learning-based architectures could attest the proposed approach's performance.

Keywords COVID-19 · Deep learning · Efficient-net · Transfer learning · Radiography imaging · Chest X-ray

✉ Hiba Mzoughi
hiba.mzoughi@yahoo.fr

¹ Advanced Technologies for Medecine and Signal (ATMS), National Engineering School of Sfax (ENIS), Route de la Soukra km 4 - 3038 Sfax, Sfax university, Sfax, Tunisia

² Gabes university, Higher Institute of Computer Science and Multimedia of Gabes, Teboulbou, Tunisia

³ Sfax university, National School of Electronics and Telecommunications of Sfax, Sfax, Tunisia

1 Introduction

By early spring 2020, a novel disease called Corona's virus (COVID-19) had been reported in most world countries. It has a disastrous effect on the global population's well-being and health. On March 11, 2020, the World Health Organization (WHO) officially classified the COVID-19 as a pandemic [19]. Conforming to data published by the WHO, by January 1, 2022, beside 300 million peoples have been infected with the COVID-19 and more than 5.4 million people have been died [17].

Effective and fast COVID-19 exploration could be as considered as an important challenge to fight fast virus transmission. The Polymerase Chain Reaction (PCR) test is the most commonly used technique for COVID-19 exploration. Although PCR testing could be assimilated as a "gold standard" used to detect the infected patients, several issues could be still about it. In the beginning, its detection's sensibility is so variable, which could produce a large amount of False Positive (FP) and False Negative (FN) results. On the other hand, due to the quick virus' spread, the poor and the heavily infected countries suffered from the lack of PCR resources "reagent kits" to ensure the enormous exploration demand.

Radiography chest exploration could be considered as an alternative solution for fast COVID-19 cases' exploration. Radiologists analyze and interpret the CT images and the chest X-ray to observe the abnormal features in order to identify the infected patients. However, for pulmonary diseases, the CT scans may have a greater sensitivity. However, one could distinguish many limitations for its practical uses in COVID-19's clinical routines exploration including the non-portability, the time scanning which could be described as long, and the exposing of the hospital staff's risk. Compared to the CT scans, the X-ray technique could be considered as:

- **Portable:** intend that scanning could be realized within an isolated room, thus could considerably reduce the risk of the virus transmittance during transport.
- **Fast:** X-Ray imaging enables rapid scanning of suspected patients and could be realized at the same time as the viral testing, which could take a long time to identify the most affected areas.
- **Available and accessible:** X-Ray imaging is available and accessible to most healthcare systems since it is considered as a standard equipment in many imaging centers. In developing countries, compared to the CT imaging which could be considered very expensive, CXR imaging is extremely readily.

However, among the biggest bottlenecks faced during the diagnosis of COVID-19 infection, are:

(1) Expert radiologists' need to explore manually the acquired radiography images due to the complex morphological features and the appearance of the chest exploration; (2) The limited number of sub-specialty radiologists, especially in the developing countries; (3) Indeed, this manual interpretation is considered as time-consuming and the diagnosis' accuracy depends on the radiologist's professional knowledge and expertise.

For these reasons, Artificial Intelligence (AI) based-aided diagnostic system is therefore highly demanded to help radiologists to detect rapidly and accurately the COVID-19 infected patients'.

Deep Learning could be defined as an Artificial Intelligence subset based on architecture inspired by the human brain ‘s function and structure. They are called Artificial Neural Networks (ANN). Especially, Convolutional Neural Networks (CNNs) have been nowadays drubbing human expertise. They are considered, actually, as a natural candidate for the processing and the analysis of various computer vision tasks.

Convolutional Neural Networks (CNNs) have been already explored to detect and to classify the chest radiography scans. Nevertheless, the challenge consists in finding a model with high accuracy and a reduced computational time to explore more scans fatly and accurately.

Recently, a study presented a new standard architecture and extended it up in order to get a novel models family, named Efficient-Nets [23]. The Efficient-Nets could be described as a deep Artificial Neural Networks subset that offers higher accuracy compared to the state-of-the-art Convolutional Neural Networks (ConvNets, “ResNet-101, ResNet-50 [8], Inception-V2 [21], Xception [5], Inception-ResNet-v2, and DenseNet-201 [10]). As revealed in Fig. 1, it reached 84.4% top-1 and 97.1% top-5 accuracy on Image-Net and it is 8.4 timeslesser and 6.1 times speedy on inferences compared to the performed ConvNet. For instance, in comparison with the commonly adopted ResNet-50, it employs analogous FLOPS, while increasing the accuracy from 76.3% to 82.6% (+6.3%).

Motivated by the high performance provided by Efficient-Net network models in deep learning applications and the urgent need to design an effective accurate and automatic solution to assist radiologists in detecting patients infected with Corona’s virus, we present, in this paper, a fully automatic, efficient and low computational classification approach to explore COVID-19 through chest X-ray images. The main contributions of this work could be summarized as follows:

- We investigate a transfer learning to the new Efficient-Nets models [23] (Efficient-NetB0-, Efficient-NetB0, Efficient-NetB1, Efficient-NetB2, and Efficient-NetB3) pre-trained on the ImageNet dataset [6] which are to the best of our knowledge rarely explored

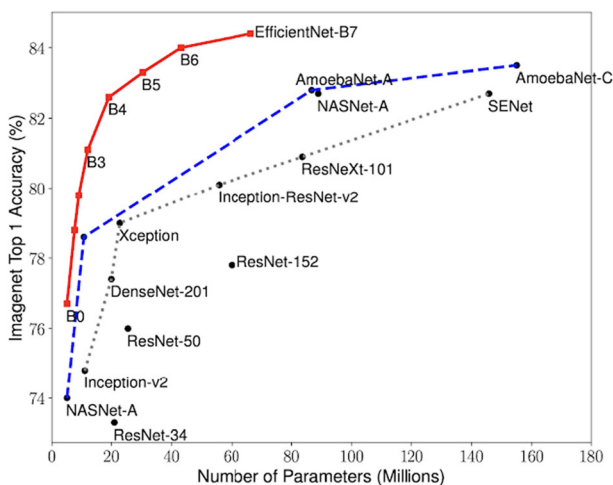


Fig. 1 Performance evaluation of Efficient-Net over the classical ImagesNet dataset [23]

- for the COVID-19 detection and classification process via radiology X-ray images. Quantitative results revealed that the best performance over the tested Efficient-Nets family models was obtained by the Efficient-Net B3 version.
- We optimize the new Efficient-net network performance for the COVID-19 detection task using the Fine-Tuning technique to select the optimal/best hyper-parameterizations of the network.
 - In order to boost the COVID-19 classification's accuracy, we balance the size of the training dataset' classes by exploiting the data augmentation techniques. Balancing the dataset classes was found to be a very effective technique and not usual in transfer learning-based methods to improve the classification results and to reduce the overfitting as well as the high variance problems.
 - We exploit, in the preprocessing task, a Min-Max normalization technique to reduce the high heterogeneity issue of radiology x-ray images since the COVID-19 training open-source dataset is usually gathered from a different source and /or sites.
 - The comparative study over two publicly available datasets attests that the proposed approach method provides high accuracy and low footprints in the same time outperforming state-of-the-art deep-learning methods such as the popular “ResNet-50 [2]”, VGG [18], and “COVID-NET [25]” achieving an overall accuracy respectively of 99% and 98% for the training and testing set.

The remainder of this paper is organized as follows: Section 2 presents a review of related works. Section 3 details the proposed approach. Section 4 explores the experimental results and discusses the obtained results. Finally, conclusions are drawn in Section 5.

2 Related works

Several Artificial Intelligence (AI) architectures based on Deep Learning (DL) have been already explored in the literature for COVID-19 prediction and classification through the radiography scan especially we could mention the pre-trained models based on convolutional neural networks architecture as the:

- **VGG-19:** Visual Geometry Group Network (VGG) is a deep neural network architecture built with three convolutional layers with several filters stacked on top of each other to improve image extraction functionality. VGGNet-19 differs from other VGG variants by adding one convolutional layer in each of the three convolutional blocks, making VGGNet-19 deeper than VGGNet-16.
- **ResNet-50:** Residual Networks (ResNet) is a deep neural network architecture. It revolutionized the CNN architectural race by introducing residual block design and selective long-range connectivity. The models stood first place in the ImageNet competition. It has a moderate architecture built with Fifth Convolutional Layer with pooling applied between the layer, and a flattened, dropout, and single fully connected layer.

Authors in [15] have proposed an automatic deep learning Computer-Aided-Diagnosis system using the chest X-ray radiology imaging to classify patients into three classes: normal (healthy), infected patient (COVID-19) and pneumonia subjects. Authors exploit several CNN pre-trained architectures such as the ResNet50, the ResNet101, the ResNet152, the

InceptionV3, and the Inception-ResNetV2 using a training dataset collected from different open sources. The pre-trained ResNet50 architecture yields the best classification rate.

The researchers in [25] have proposed the first open-sourced deep neural network called COVID-NET for COVID-19 case exploration through CXR scans. The authors report an overall accuracy and overall sensitivity of 92.4% and 80% respectively. In-Despite their success, the COVID-NET is considered as a perplexing network, which claims 2.1GB of memory and achieves more than 3.5 billion MAC operations denoting three principal drawbacks: the time computation, the consumption cost, and the execution time of these models remaining very high, which could inflict many difficulties to their availability and accessibility. After that, they evaluated and compared the performances of three state-of-the-art pre-trained deep learning models such as the VGG [18], the ResNet50 [21], and the Inception-V3 [8] in order to classify CXR scans into COVID-19 class and non-COVID-19 class. The training data is augmented using the random rotation; the horizontal flips, the data augmentation techniques, and the random noise. The experiment results reveal that the VGG-19 and the Inception provide the best-performing architectures.

Authors in [4] have proposed a multi-classification method into three classes (pneumonia, safe, and infected with COVID-19) using the CT scans. The authors investigated transfer learning to the convolutional blocks of four versions of the popular ResNet using the public image collections with a limited training dataset. The higher scores reached by the pre-trained ResNet50 architecture for COVID-19 versus all other groups are 91.21% accuracy, 88.46% specificity, and 94.87% sensitivity. The global accuracy of the three-category classification over the testing set [22] is 89.01%. The authors have collected a large dataset with 8009 CXR scans.

Three CNN based-architectures were trained using the collected dataset: the Faster R-CNN [7], the Mask R-CNN [26], and the ResNet-50. Best accuracy was obtained using the Faster R-CNN (87%). The Mask R-CNN accuracy attains 83% but the ResNet-50' accuracy is only 72%. The Error framework has been adopted to evaluate the error quantity. The main drawback is that the training takes much time as it was performed for several hours as mentioned by the authors.

Study [24] is based on the open-source COVID-NET [25] to build a new model called "EDL-COVID". To reduce the computation cost and the training's required time, the authors adopt an alternative solution based on generating a cascade of multiple deep learning architectures and executing only one training run to predict the final results. However, this solution has a big problem which is producing similar models. In fact, using analogous models aims to obtain comparable predictions and prediction errors, meaning that this model's combination could not provide many benefits.

On the other hand, researchers in [11] have introduced a new hybrid model in which the deep features vector was excerpted from the ResNet50 architecture and fed to the linear kernel function of an SVM classifier. The CXR dataset including 200 healthy and 180 infected cases with COVID-19 chest X-ray scans has been used in the experimentation study. The Classification accuracy measurement has been adopted to evaluate the approach's efficiency. The experimental results revealed good classification accuracy which reached 94.7%.

Meanwhile, the conventional models pretend to grant sufficient results, the proposed deep learning models' reliability could be questioned due to a crucial bias problem for the COVID-19 database regrouped from a reduced COVID-19 cases' sized group. Hence, the deep learning model's prediction results are exposed to undergo due to the generalization errors and the high

variance. Such problems could result from the limited amount of the database and the noise. Moreover, the training processes relying on stochastic approaches could be sensitive and could even generate different weights at each training.

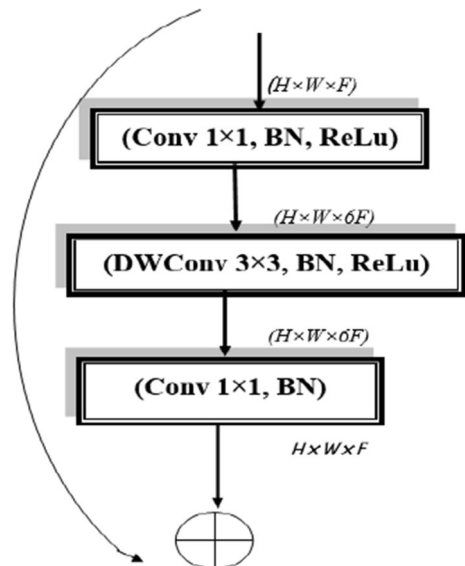
To alleviate these issues, we propose, in this paper, to adopt the transfer learning technique on a pre-trained Efficient-nets model. Efficient-Nets families (B0 to B7) are characterized by high accuracy and low footprints. The architecture for the baseline network “EfficientNet-B0” is based on the principal building component of the Mobile Inverted Bottleneck Conv (MBConv) Block [16]. The MBConv architecture is illustrated in Fig. 2.

The MBConv’ blocks building by pass between the starting and the ending of each convolutional block. The activation maps’ input is extended by adding (1×1) convolution layer to improve the feature maps’ depth. In this model, the shortcut connections rely on the narrow layers. At the same time, the wider layers are placed in the middle of the jump links. The considered model aims to decrease the model’s size as well as the required total number’s transactions.

3 Proposed method

Deep learning-based models could be considered as complex, and therefore the required memory and the computation time could be important. On the other hand, several trainable parameters are suggested to avoid overfitting. Indeed, one of the serious problems during training deep neural networks is that the dataset is not abundant. Therefore, the learned network has a better performance on the training set meanwhile has an under-performance on the test set. To alleviate these issues, we adopt two techniques: (1) transfer learning and (2) data augmentation techniques. Fig. 3 illustrates the proposed flowchart of the deep learning-based COVID-19 diagnosis approach.

Fig. 2 MBConv Block. DWConv stands for depth wise conv, $k3 \times 3/k5 \times 5$ defines the kernel size, BN is batch norm, $H \times W \times F$ means tensor shape (height, width, depth) [13]



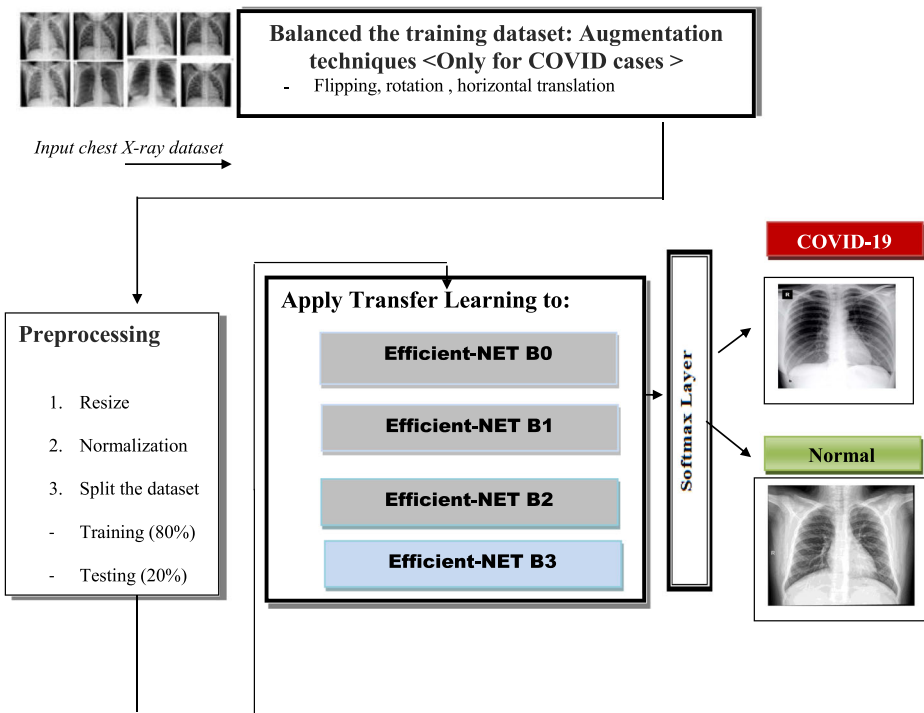


Fig. 3 The proposed flowchart of deep learning based COVID-19 diagnosis approach

We employed transfer learning to Efficient-Nets models family pre-trained using the ImageNet dataset. The classical ImageNet dataset is much broader and provides an abundance of training data. We could, then, obtain over than 14 million scans. Therefore, the carried network weights could be taken initially and could be all **fine-tuned** using the Adam optimizer during the following training step. By permitting the fine-tuned of all the weights, we enable the architecture to be specific to the considered problem.

The proposed approach steps for accurate detection of COVID-19 could be summarized as follows:

1. **Dataset Balancing:** in favor of the minatory classes (COVID-19 cases) in a manner that the images' number is equitable for the two classes after the transformation's application. The techniques used for dataset augmentation are flipping, random rotation, and horizontal translation.
2. **Preprocessing:** our proposed preprocessing consists of two steps: image intensity normalization and resizing.
3. **Transfer learning using Efficient-Net models:** One could use the pre-trained network's weight to reduce the training time. For this reason, transfer learning is applied to Efficient-Nets pre-trained using the ImageNet dataset. Then, we explore and test the performance of four Efficient-Nets versions: Efficient-NetB0, Efficient-NetB1, Efficient-NetB2, and Efficient-Net B3. The best performance has been obtained using Efficient-Net B3; we fine-tuning it to select the optimal network hyper-parameters.

4. **The final classification:** we classify the dataset into two classes: normal cases and infected patients (COVID-19).

The proposed model has been trained and tested. The accuracy metrics has been evaluated to improve the efficiency. In this part, we are going to detail the different steps of the proposed approach:

- *Dataset augmentation: the Problem of imbalanced training dataset classes:*

In most of deep convolutional learning algorithms, the classification task with unequal classes' distribution has undergone an important drawback in terms of performance. For this reason, it is better to ensure balanced data for considered classes.

A study [20] was conducted to prospect the relationship between the classification accuracy and the training data's class distribution. It demonstrates that better results could be obtained when considering balanced distribution since the neural network learning process aims to boost the majoritarian classes' classification.

In this work, we found that our training dataset is significantly with imbalanced classes' distribution. To alleviate this problem, we up-sampled the training dataset in such a way that the images' number in both classes is equal after the transformation application.

Data augmentation aims to increase the training dataset by applying some transformations [32]. It should ensure that the semantic information is conserved. The techniques adopted for dataset augmentation, in this work, are shown in Table 1.

3.1 Data preprocessing

- *Image Resizing*

In the present study, the initial size of images is (512×512) and was resized to (224×224) . We used the Compose method (from Torchvision module in Pytorch libraries) to resize the images.

- *Intensity Normalization*

In medical radiology imaging analysis, the training images are usually gathered from different sites and sources, therefore, the acquired images' intensities profile could be considered highly heterogeneous. For this reason, we applied an intensity normalization technique that relied on linear transformation in the range $[0, 1]$ using the Min-Max normalization technique.

- *Dataset Splitting:*

Table 1 Used dataset augmentation techniques

Data augmentation technique	# Parameters	Description
Rotation	10	Apply specified angle or degree of rotation to the image
Horizontal Flip	0.7	Horizontally reversing the entire rows and columns of an image pixels

The used dataset is divided as follow: 80% for the training process and 20% for the testing task. In fact, one of the effective techniques, besides to data augmentation, found to be effective to prevent /avoid overfitting in DL based-approaches is to use the majority of data for the training process.

3.2 Transfer learning using efficient-net models

Typically, in deep-learning, when designing a deep network architecture from scratch, a large scale dataset is needed during the training process to obtain the desired' network performance and to alleviate the overfitting effect. However, when adopting a transfer learning technique, we could exploit and transfer pre-trained weights directly from layers using a large scale and much border dataset such as the Image-Net dataset to enhance the generalization and the learning function of a new different task. This technique has proved an efficient performance in term of improving the overall network' accuracy and accelerating the training time even when pre-trained weights are transferred from a totally different task.

In the proposed approach, we applied transfer learning to Efficient-Nets family models pre-trained on the ImageNet dataset [31] which is obviously much border dataset then the used chest x-ray radiology images. Usually, deep convolutional models try to capture the contour and the edges features during the learning process from the initial layers of the network, the shape are captured from the middles layers, and relevant features are captured from the last layers. In transfer learning, the features captured from the first and the middle layers are maintained and can be re-used in many tasks. The pre-trained weights are used as an initial solution and are next fine-tuned using the adopted optimizer. The concept about is that the pre-trained models have gained a lot of knowledge about different objects. Therefore, by allowing all the weights to get fine-tuned, we are permitting the model to specialize the specific task. The different steps of the transfer learning process could be summarized as follows:

1. Duplicating the obtained weights from the first model (pre-trained) to the novel deployed architecture;
2. Designing the new model's design to prepare it to the new problem where new layers could be included.
3. Initializing the added layers;
4. Using a suitable optimization algorithm, the training could be started by updating the weights conforming to the loss function.

– *Efficient-Net network architecture*

The Efficient-Net is defined on the baseline of deep artificial neural networks. In contrast to the state-of-the-art approaches that randomly adjust the network dimensions (width, depth, resolution ...), the Efficient-Net models systematically adjust each of its dimensions using scaling coefficients' fixed set. Each dimension is scaled by the parameter according to the following equations:

$$\text{Depth} \propto \alpha^{\varnothing} \quad (1)$$

$$Width \quad \beta^\emptyset \tag{2}$$

$$\begin{aligned} Resolution \quad \gamma^\emptyset \\ \text{s.t. } \alpha, \beta^\emptyset, \gamma^\emptyset = 2 \\ \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{aligned} \tag{3}$$

Where α , β and γ are constants collected by a framework search experiment. Instinctively, \emptyset is a consumer particularized-specified coefficient which manages the available resources for the model adjustment. The variables α , β and γ designate how to select extra resources such as: the width, the depth, and the resolution respectively. Figure 4 presents the Efficient-Net baseline network architecture.

For our experimentation study, we used the Efficient-NetB3. Its architecture is illustrated in Table 2. It reduces flops by adopting a multi neural search model. The Efficient-Net’s convolution layers could be classed into two sections: the point-wise convolution and the depth-wise convolution. Such process is intended to reduce the calculation time, improve accuracy, and reduce Loss.

The Efficient-Net’ layers have been raised by maintaining a fixed constant ratio to boost the architecture accuracy and performance. The last layer’s output is fed to the convolutional layer with an “average pooling” down-sampling strategy by a 3×3 kernel size. The output is next feed to a flatten to convert the features matrix into vectors which are next passed a Fully Connected (FC) layer with the “ReLU” activation function. In the final layer, “Softmax” activation is employed to classify the images into two classes.

- **Presentation of network hyper-parameters**

Fine-tuning To ensure the best network performance and to select the optimal network’ hyper-parameters, all the pre-trained networks are fine tuned.

– **Number of Epoch**

Numerous epochs are tested, and one could notice that for the considered database, with the pre-trained model, the best test accuracy has been attained after 11 epochs.

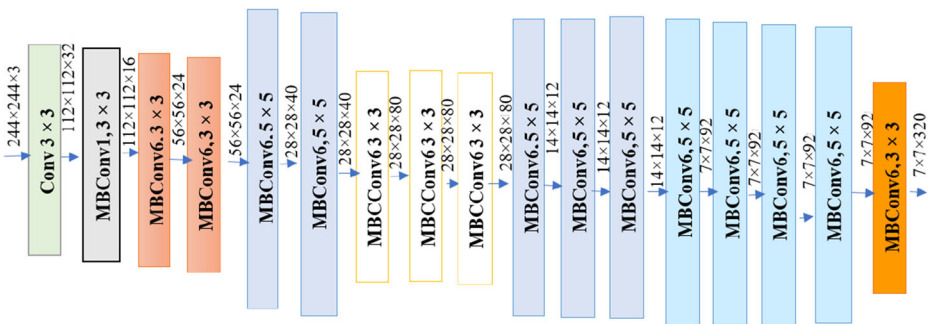


Fig. 4 Efficient-Net baseline network architecture.: The efficient-Net B0 [1]

Table 2 The Efficient Net B3 architecture Conv refers to convolutional layer, FC refers to Fully Connected (FC) Layers

Stage	Operator	# Number Layers	Activation function	Resolution	#Number of Channels
1	Conv,kernel seize=(3×3)	1	ReLU	244×244	32
2	MBCConv,kernel seize=(3×3)	1	Relu	112×112	16
3	MBCConv,,kernel seize=3×3	2	Relu	112×112	24
4	MBCConv,,kernel seize=(5×5)	2	Relu	56×56	40
5	MBCConv,,kernel seize=(3×3)	3	Relu	28×28	80
6	MBCConv,,kernel seize=(5×5)	3	Relu	14×14	112
7	MBCConv,,kernel seize=(5×5)	4	Relu	14×14	192
8	MBCConv,k5×5	1	Relu	7×7	320
9	Conv1x1/Pooling/FC	1	Softmax	7×7	1280

– Loss Function

The cross-entropy has been adopted to calculate the loss since we are handling a classification task. The loss functions are used to minimize the prediction error.

– Optimizer

To make our predictions as correct and optimized as possible, three different optimizers have been tested:

1. **RMSprop**: In neural network training, RMSprop is a gradient-based optimization strategy. This normalization equalizes the step size (momentum), lowering it for high gradients to avoid exploding and raising it for minor gradients to avoid vanishing problems.
2. **Adam**: could be considered as a combination of Stochastic Gradient Descent and RMSprop, with the momentum's addition. Similarly, to the RMSprop, it adjusts the learning rate using the squared gradients, and it receives the momentum's advantage.
3. **SGD optimizer**: First-order methods are currently dominating neural network training in the stochastic gradient descent.

– Activation function

The Rectifier linear unit (ReLU) is responsible of the nonlinearity which converts the data. ReLU have been adopted in this study as an activation function. It could be defined by eq. (4):

$$f(i) = \text{Max}(0, i) \quad (4)$$

where $f(i)$ represents the function of neuron's output of an input called "i".

4 Experimental results

The experimental results and the implementation setup were detailed in this section.

- **Implementation Details**

The execution environment of the computational experiments was conducted on an Intel(R) Core(TM) i7-5820K CPU @ 3.30GHz, equipped with NVIDIA GPU Geforce GTX of 64 GB Ram. The proposed approach was built using Pytorch Library which is a Deep Learning /Machine Learning framework for Python development environment suitable for Computer Vision tasks. Python 3.7 development environment was used to construct the proposed model.

- **Dataset description**

In order to validate and to evaluate the performance of the proposed approach, two chest X-Ray radiology imaging dataset has been used.

- Dataset 1

Recently, researchers from the University of Waterloo and the University of DarwinAI Corp., both from Waterloo, Canada collected the largest open-access benchmark dataset available in terms of the number of publicly available COVID-19 positive cases. The COVIDx dataset is collected from five different data sources. The dataset selection used for the training and the testing is a publicly available challenge [12] in order to detect and classify normal and COVID-19 cases.

The training dataset is a supervised dataset regrouping 15,264 images that have been classified by an expert. However, the test dataset is composed of 400 such images. Figure 5 illustrates an example of positive and negative cases from the tanning dataset.

- Dataset 2

Recently researchers published an open-source X-Ray radiology imaging dataset in the GitHub repository [3]. This database has been used to detect and to classify COVID-19 cases. The supervised dataset includes 100 Positive (infected with COVID-19) patients and 100 Negative (Normal) chest X-Ray radiology images. The size of the dataset increased significantly after applying our proposed augmentation techniques. Figure 6 donates two examples of positive and negative subject cases from the training set.

- **Performance Evaluation Metrics**

Besides to accuracy measurements, the sensitivity, specificity, precision, **Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F-score metrics** have been used to evaluate the efficiency and the productivity of each pre-trained model tested.

True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) are the four expected results, which are required to calculate the assessment criteria.

- *TP*: has been classed as COVID-19 infected, which is correct.
- *TN*: stands for “not infected with COVID-19,” which is correct.

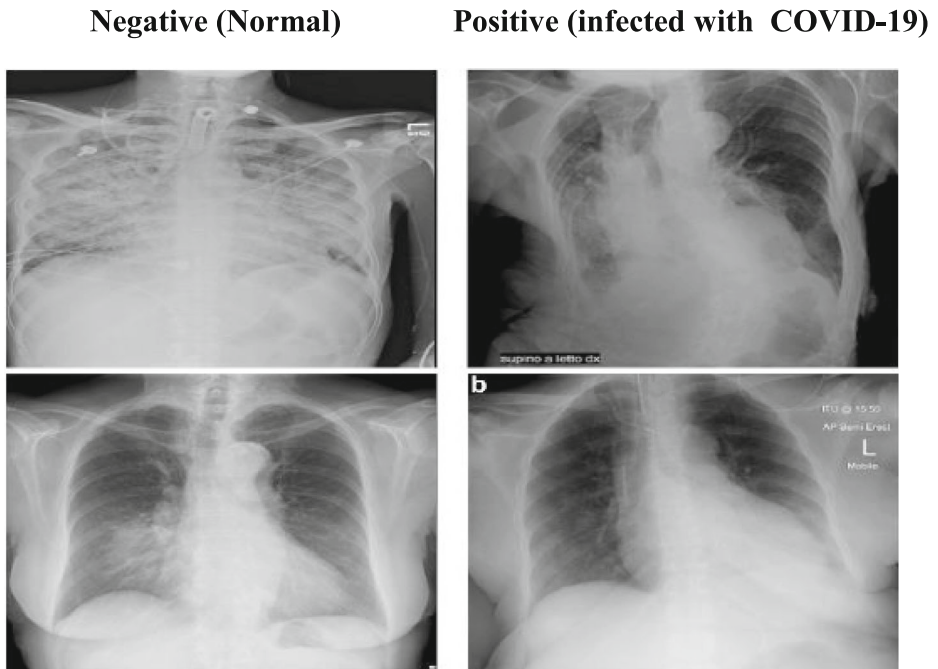


Fig. 5 Examples of Positive and Negative cases from training dataset 1

- *FP*: classed as COVID-19 infected, although this is incorrect.
- *FN*: incorrectly classed as NOT infected with COVID-19.
- **Accuracy**: As shown in Eq. (5), accuracy is equal to the sum of true positives and true negatives divided by the total values of the confusion matrix components.

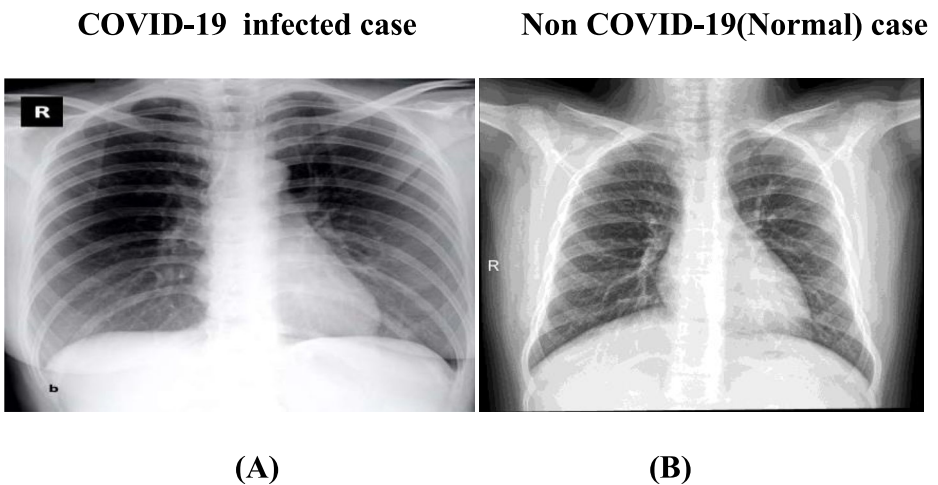


Fig. 6 Example of positive and negative sujet cases from training dataset 2 (A) infeted sujet-case with COVID-19 image (B) normal subject cases

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (5)$$

- **Sensitivity:** the *sensitivity* could be calculated as the ratio between the proportion of the positive infected patients that has been successfully predicted and the total number of samples Eq. bellows illustrates the formula of sensitivity metric.

$$\text{Sensitivity} = \frac{TP}{FN + TP} \quad (6)$$

- **Specificity:** the *specificity* represents the ratio between the normal subject-cases (or samples) that are predictive as negative samples and the total number of samples. Specificity is calculating according to Eq. (7)

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (7)$$

- **Positive Predictive Value (PPV):** represents the probability of positive results that are predicted as truly positive infected subject-cases. PPV measurement could be calculated according to Eq. 8

$$\text{PPV} = \frac{TP}{TP + FP} \quad (8)$$

- **Negative Predictive Value (NPV):** represents the probability of negative results that are predicted as negative subject-cases. NPV metric could be calculated according to Eq. 9

$$\text{NPV} = \frac{TN}{FN + TN} \quad (9)$$

- **Positive Likelihood Ratio (LR+)** could be expressed as the ratio between the probability that a subject-case is predicted as positively infected (true positive) and the probability of negative subject case that is predicted as positive (false positive). It is calculated according to Eq. 10

$$\text{LR+} = \frac{\text{Sensitivity}}{1 - \text{Specificity}} \quad (10)$$

- **Negative Likelihood Ratio (LR-)** could be expressed as the ratio between the probability of positive/infected subject-case is predicted as negative (false negative) and the probability of negative subject case that is predicted as negative (true negative). It is calculated according to Eq. 11

$$\text{LR-} = \frac{1 - \text{Sensitivity}}{\text{Specificity}} \quad (11)$$

- **Precision:** the precision represents the difference between the number of samples that are predicted as true positive and the total number of samples that are projected to be positive. Eq. 12 shows precision mathematical formula.

$$Precision = \frac{TP}{FP + TP} \tag{12}$$

- **F-Score:** The F1 score represents the harmonic mean of the precision metric and the sensitivity metric. It could be calculated according to Eq. (13):

$$F-score = \frac{2TP}{2TP + FP + FN} \tag{13}$$

- **Area Under Receiver Operating Characteristic (AU-ROC) Curve;** is exploited to evaluate a binary classification method. It is plotted in function of the False Positive Rate (FPR) and the True Positive Rate (TPR).

To find the performed model among the Efficient-Nets family, four architectures (versions) are tested: The Efficient-Net B0, the Efficient-Net B1, the Efficient-Net B2, and the Efficient-Net B3. Tables 3 evaluates the performance of the tested Efficient Net models that are scaled from the same baseline EfficientNet-B0. We could notice that all the networks have similar accuracy’s performances using the training dataset. However, among the presented models, the EfficientNet-B0 is the simplest one.

The results, shown in Table 3, confirm that the proposed approach was able to produce a high-quality model, with an overall accuracy of 99%, for both classes. The training accuracy, in terms of the number of epochs for different tested Efficient-Net models, is presented in Fig. 7.

– **AU-ROC Curve**

To assess the performance of the proposed method, the AU-ROC curve has been realized. Figure bellow shows the graphical plot of the AU-ROC curve of our proposed method Fig. 8.

Table 4 provides the performance evaluation over the testing dataset. According to experimental results, the best testing accuracy is obtained using the Efficient-Net B3 model.

The Table below (Table 5) illustrates the classification’s results of COVID-19 subject cases using two publicly available datasets. We could notice that the proposed approach presents a comparable performance for the both explored datasets using the Efficient-Net B3 model for the training sets.

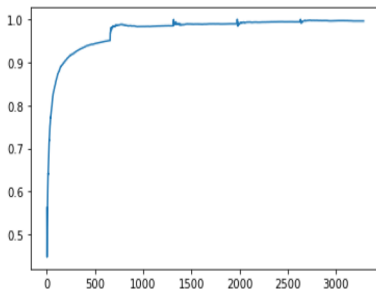
According to the evaluation’s performance using accuracy measurement when considering the training and the testing sets of the two datasets, superior performance is obtained for the testing set when dataset 1 is used, this could be explained that dataset 1 is much larger data when compared to the second data.

Table 3 Performance evaluation of the proposed models using the training images

The model	# Pram	Input-shape	Accuracy of 0	Accuracy: of 01
Efficient-Net B0	5.3 M	(244,244)	Accuracy of 0: 99%	Accuracy of 1: 99%
Efficient-Net B1	7.8 M	(244,244)	Accuracy of 0: 99%	Accuracy of 1: 99%
Efficient-Net B2	9.2 M	(244,244)	Accuracy of 0: 99%	Accuracy of 1: 99%
Efficient-NET B3	12 M	(244,244)	Accuracy of 0: 99%	Accuracy of 1: 99%

Training dataset accuracy using the *Efficient-Net*

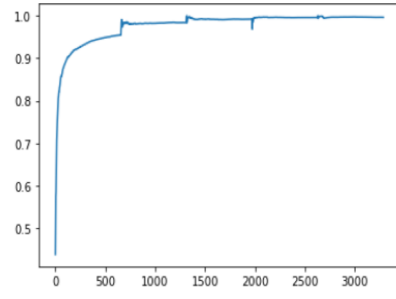
B0 in term of number of epochs



(a)

Training dataset accuracy using the *Efficient-*

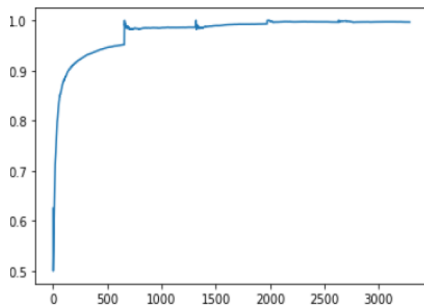
Net B1 in term of number of epochs



(b)

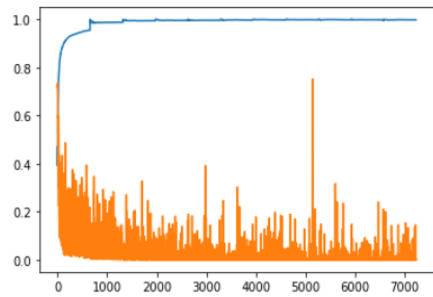
Training dataset accuracy using the *Efficient-Net*

B2 in term of number of



(c)

Training dataset accuracy (blue) and Loss (orange) using *Efficient-Net B3* model



(d)

Fig. 7 (a) Training dataset accuracy using the Efficient-Net B0 in terms of number of epoch; (b) Training dataset accuracy using the Efficient-Net B0 in terms of number of epoch; (c) Training dataset accuracy using the Efficient-Net B2 in terms of number of the epoch; (d) Training dataset accuracy (blue) and Loss (orange) using efficient-Net B3 model

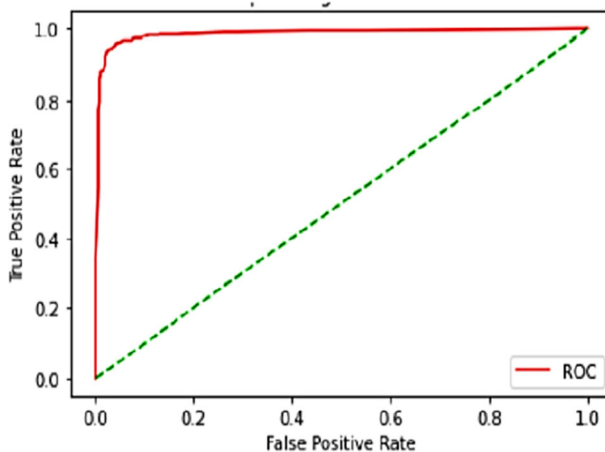


Fig. 8 AU-ROC curve of the proposed approach

Table 4 performance evaluation over the Testing dataset accuracy

The model	# Pram	Input-shape	Overall Testing Accuracy (%)
Efficient-Net B0	5.3 M	(244,244)	0.988
Efficient-Net B1	7.8 M	(244,244)	0.987
Efficient-Net B2	9.2 M	(244,244)	0.977
Efficient-Net B3	12 M	(244,244)	0.990

– Optimizer

The Stochastic Gradient Descent (SGD), the RMSprop and the Adam optimizers have been proved to judge the classification efficiency. The initial learning rate for both optimizers was set to 0.0001 during the experiments. Table 6 represents the parameters of tested optimizers.

Fig. 9 provides a performance evaluation of the tested optimizers. According to the obtained results, the “Adam” and the “RMSprop” optimizers presents comparable performance. Meanwhile; SGD presents the lowest accuracy measurement. As illustrated in Fig. 9 the Adam optimizer provides the best performance (99%) for both training and testing dataset. This scenario is adopted to judge the considered network.

– Activation function

In order to select the optimal hyper-parameters of the network, two activation functions have been tested mainly the “Softmax” and the “Sigmoid”. Fig. 10 presented a performance evaluation of the studied activation functions in term of classification overall accuracy using testing and training sets.

We noticed that best network performance was obtained using the Softmax activation function by achieving an overall accuracy of 0, 99% using the testing set, meanwhile the Sigmoid has achieved a lower performance (0,8125) on the testing set. Based on the performance evaluation, Softmax has been adopted as an activation function for the proposed approach.

– Impact of the dataset balancing

To assess the impact of balancing the size of the dataset, we measured the overall classification’s accuracy first without applying any data augmentation techniques and next using dataset with the same classes’ distribution. The results are illustrated in Table 7.

One could notice that the accuracy is clearly better when using the same distribution of the dataset; the classification’s accuracy enhances from 0.86 to 0.990.

– Impact of dataset splitting

Several work studied the impact of dataset splitting on the classification accuracy [14]. The dataset splitting ratio could affect the classification’ result. To assess the impact of dataset

Table 5 Accuracy measurement of the proposed approach using two datasets over training and testing sets

	Overall Training accuracy	Overall Testing accuracy	Input-shape	Efficient-Net B3 (param)
Dataset 1 [12]	0,99	0,99	(244,244)	12 M
Dataset 2 [3]	0,99	0.8750	(244,244)	12 M

Table 6 Parameters of tested optimizers

Optimizer	Parameters
Adam	- Learning rate=1e-4 - Beta_1 (decay rate for the first moment estimates)=0.99 - Beta_2 (decay rate for the 2nd moment estimates) =0.99 - Epsilon=1e-07
RMSprop	- Learning rate=1e-4 - Rho=0.99 - Momentum=0.9 - Epsilon=1e-07
SDG	- Learning rate=1e-4 - Momentum=0.9

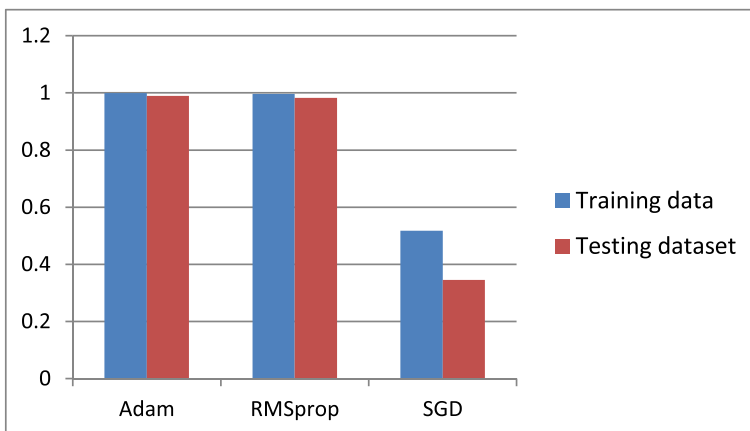
splitting technique on the classification' accuracy, we randomly split our dataset into training set and testing set. Results are summarized Table 8.

According to results in Table 8, the best accuracy was obtained using 80% for training set and 20% for testing set.

5 Comparative study and discussion

Table 9 presents the proposed approach performance using various evaluation metrics (as sensitivity, specificity, precision, PPV/NPV, Positive and Negative Likelihood Ratio and the F1-Score) as well as a comparative study with some existing works from the state-of-the-art for COVID-19 detection.

We compare our approach with the COVID-NET architecture proposed recently by Wang et al. [25] besides to a comparison with studies based on transfer learning of pre-trained models based on convolutional networks such as the “ResNet-50” and “VGG-16” and the “VGG-19” version models. In fact, the choice of these three architectures is based on the fact that they are among the most popular architectures explored largely in the literature for COVID-19 detection task. Furthermore, we compared our work with the Mobile-Net models which are recently investigated in the literature and provide a good performance with low cost.

**Fig. 9** comparative results in term of accuracy using of different tested optimizers

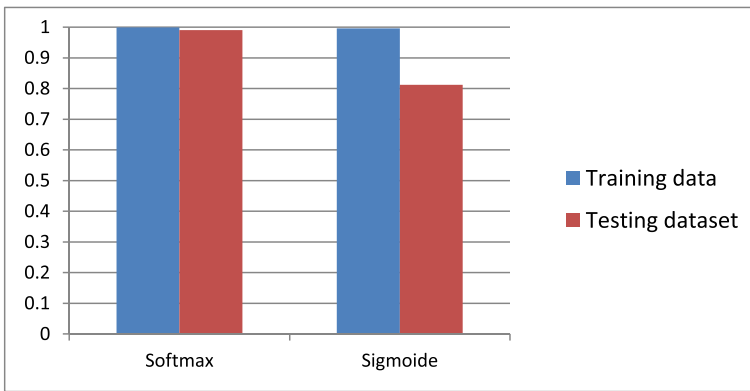


Fig. 10 comparative results in term of accuracy using of two activation functions

In fact, the cost of a model depends on the number of generated parameters. Models that generated a huge number of parameters require a huge amount of data to adjust them. Hence, the major bottleneck faced by complex model in the classification task of chest X-Ray radiology images into (normal/infected) cases is the need of a bigger dataset. That may explain the lower performance of very complex models especially for the minatory classes of the dataset.

As it obvious in Table 9, that the compared methods have a similar performance. However, the more complex model is, the lower performance is given. For instance, the VGG-16 is a deep model which generates more than 138,357 parameters that stacks deep convolutional layers with small (3 × 3) kernels size, max-pooling along a stride of 2 and two Fully Connected (FCs) layers followed by Softmax activation layer. With the increasing dept. of the network, the models become slower during training. The reported results were about 77.0%, 67.7% and 63.64% for the accuracy, the sensitivity and the precision respectively. The other metrics was not reported in the compared works.

COVID-NET is considered as a complex model regarding the huge generated number of parameters. The proposed approach has presented an enhancement of the number of infected patients that has been successfully predicted (98.02%) compared to COVID -NET model (91.0%) according to the sensitivity measurement. Furthermore, a significant reduction of

Table 7 the impact of dataset balancing in term of accuracy

	Wwithoutt balanced	With balancing
Training set	0.94	0,998
Testing set	0.86	0,990

Table 8 Impact of dataset splitting on the accuracy

Training set (%)	Testing set (%)	Training Accuracy	Test accuracy
80%	20%	0,99	0,99
70%	30%	0,9500	0,8778
60%	40%	0,93	0,78

Table 9 Comparative study of the proposed method with other existing works bases on the testing dataset 1

Architecture model	Accuracy	Sensitivity	Specificity	PPV	NPV	positive and negative likelihood ratio	F1-Score	Precision	#Params
VGG-19 [18]	83.0%	58.7	-	-	-	-	-	-	143,667,240
ResNet-50 [8]	90.6%	83.0	-	-	-	-	-	-	25,636,712
COVID-NET [25]	93.3%	91.0	-	-	-	-	-	-	126.6 GB
MobileNet [13]	90.4%	83.8%	-	-	-	-	-	100.0%	4,253,864
Mobile Net V2 [13]	90.0%	87.1%	-	-	-	-	-	96.4%	3,538,984
VGG-16 [13]	77.0%	67.7%	-	-	-	-	-	63.64%	138,357,544
Proposed Approach	99.0%	98.02	99.0%	98.04%	99%	98.01% 0.010102	0.99	99.0	11.6

number of parameters was obtained with the proposed approach. When compared to COVID - NET, our proposal is characterized by a considerably lower *architectural complexity and lower computational complexity*. For example, for the smaller model; the COVID-NET requires memory of 2.1GB and achieves over 3.5 billion MAC operations involving three main drawbacks: time consumption, infrastructure cost, and computation cost. The authors investigated a data augmentation strategy in the COVID -NET implementation; however, there is no information about the new size of the data after applying augmentation techniques. Since the convolutional network tends to enhance the accuracy results of majority classes. This may justify the superiority of proposed approach obtained by balancing the size of dataset ‘classes.

Compared to VGG models, ResNet have a lower complexity since it is based on ‘*average pooling*’ rather than the ‘*fully connected layer*’. ResNet is deep network architecture rely on the residual modules/blocks using SGD optimizer. With a totality of 50 convolutional layer with different size of depth, the model’s architecture is considered as complex. Moreover, the introduction of skip connections adds a constraint to take into consideration the dimension of layers. In computer vision task, ResNet-50 achieves good performance especially in COVID - 19 detection and classification task. An overall accuracy of 90, 6% was obtained.

The Mobile-Net architecture [9] is a lightweight network architecture and computationally not exhausted compared to VGG and ResNet models. It rely on depth-wise separable convolution. Two hyper-parameters were introduced to decrease the cost and the computational requirement mainly the “*width multiplier*” refers a α and “the resolution wise multiplier” refers as ρ . In order to narrow a network, the input and the output channels of the network ‘layers will be multiplied by the width. Meanwhile, the input feature map for a given layer will be multiplied by the resolution wise parameter “ ρ ”.

Although the Mobile-Net is a simple and a computationally not exhausted, when compared with the MobileNet, one could notice that the proposed model provides a higher accuracy value on the testing dataset by achieving 99% test accuracy while providing a lower number of parameters. Thus could attest the efficiency of the proposed approach.

6 Conclusion

We presented, in this paper, an efficient and a low computational deep learning-based approach to detect COVID-19 cases using chest X-Ray images. Instead of training a model from scratch, we investigated transfer learning techniques to accelerate the learning process. Performance evaluation analysis shows that using the same distribution of dataset classes through using dataset augmentation techniques helps to boost the classification accuracy. Furthermore, meanwhile, the majority of deep learning-based approaches have superior performance on the training dataset and underperformance on the testing set; our approach has a comparable performance for both sets.

The comparative study reveals that the proposed approach outperforms state-of-the-art well-known CNN-based architectures for COVID-19 exploration. For future works, we propose to implement the current model in medical equipment or even radiologists’ mobile phones. We are further interested to include Grad-CAM evaluation of our proposed method for a more in-depth analysis as well as to extend our proposed method to distinguish between Pneumonia bacterial and COVID-19 samples in the future works.

Data availability COVID-19 training and testing dataset are available on:

<https://www.kaggle.com/c/stat946winter2021/data>

Disclosures Authors have nothing to disclose.

References

1. Ahmed T, Sabab NHN (2022) Classification and understanding of cloud structures via satellite images with EfficientUNet. *SN Comput Sci* 3:1–11
2. Akiba T, Suzuki S, Fukuda K (2017) Extremely large minibatch sgd: Training resnet-50 on imagenet in 15 minutes. *arXiv Prepr. arXiv1711.04325*
3. CGAN-Covid19-SyntheticImageGeneration_Dataset at main · tejas20_CGAN-Covid19-SyntheticImageGeneration · GitHub. (n.d.)
4. Chen H, Guo S, Hao Y, Fang Y, Fang Z, Wu W, Liu Z, Li S (2021) Auxiliary diagnosis for COVID-19 with deep transfer learning. *J Digit Imaging* 34(2):231–241. <https://doi.org/10.1007/s10278-021-00431-8>
5. Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1251–1258
6. Deng J, Dong W, Socher R, L.-J. Li, K. Li, Fei-Fei L (2009) Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp 248–255
7. Girshick R (2015) Fast r-cnn, in Proceedings of the IEEE international conference on computer vision. pp 1440–1448
8. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 770–778
9. Howard AG et al (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv Prepr. arXiv1704.04861*
10. Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K (2014) Densenet: Implementing efficient convnet descriptor pyramids. *arXiv Prepr. arXiv1404.1869*
11. Ismael AM, Şengür A (2021) Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Syst Appl* 164:114054. <https://doi.org/10.1016/j.eswa.2020.114054>
12. Kaggle, COVID-19 X-ray image classification | Kaggle (2020) [Online] Available: <https://www.kaggle.com/c/stat946winter2021>
13. Luz E, Silva P, Silva R, Silva L, Guimaraes J, Miozzo G, Moreira G, Menotti D (2022) Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images. *Res Biomed Eng* 38(1):149–162
14. Meng Z, McCreddie R, Macdonald C, Ounis I (2020) Exploring data splitting strategies for the evaluation of recommendation models. In: Fourteenth ACM conference on recommender systems. pp 681–686.
15. Narin A, Kaya C, Pamuk Z (2021) Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Anal Applic* 24:1–14
16. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 4510–4520
17. Semenzato L et al (2021) Chronic diseases, health conditions and risk of COVID-19-related hospitalization and in-hospital mortality during the first wave of the epidemic in France: a cohort study of 66 million people. *Lancet Reg Heal* 8:100158
18. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv Prepr. arXiv1409.1556*
19. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, al-Jabir A, Iosifidis C, Agha R (2020) World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). *Int J Surg* 76:71–76
20. Sun Y, Wong AKC, Kamel MS (2009) Classification of imbalanced data: a review. *Int J Pattern Recognit Artif Intell* 23(4):687–719. <https://doi.org/10.1142/S0218001409007326>
21. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 2818–2822

22. Tahir H, Iftikhar A, Mumraiz M (2021) Forecasting COVID-19 via registration slips of patients using ResNet-101 and performance analysis and comparison of prediction for COVID-19 using faster R-CNN, mask R-CNN, and ResNet-50. Proc. 2021 1st Int. Conf. Adv. Electr. Comput. Commun Sustain Technol ICAECT 2021. <https://doi.org/10.1109/ICAECT49130.2021.9392487>
23. Tan M, Le QV (2019) EfficientNet: Rethinking model scaling for convolutional neural networks. 36th Int. Conf. Mach. Learn. ICML 2019, vol. 2019-June, pp. 10691–10700
24. Tang S, Wang C, Nie J, Kumar N, Zhang Y, Xiong Z, Barnawi A (2021) EDL-COVID: ensemble deep learning for COVID-19 case detection from chest X-ray images. IEEE Trans Ind Inform 17(9):6539–6549. <https://doi.org/10.1109/TII.2021.3057683>
25. Wang L, Lin ZQ, Wong A (2020) COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci Rep 10(1):1–12. <https://doi.org/10.1038/s41598-020-76550-z>
26. Wei X-S, Xie C-W, Wu J, Shen C (2018) Mask-CNN: localizing parts and selecting descriptors for fine-grained bird species categorization. Pattern Recogn 76:704–714

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.