International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Mar-2024)

http://dx.doi.org/10.12785/ijcds/1501102

Automatic Detection of COVID-19 from Chest X-Ray Images Using EfficientNet-B7 CNN Model with Channel-wise Attention

Mohamed Rami Naidji[1] **and Zakaria Elberrichi**[1]

[1]*Computer science Department, EEDIS Laboratory, Djillali Liabes University, Sidi Bel Abbes, Algeria*
Received 2 Feb. 2024, Revised 11 Mar. 2024, Accepted 12 Mar. 2024, Published 15 Mar. 2024

Abstract: Since the outbreak of the global COVID-19 pandemic in Wuhan, China, in 2019, its impact has been seen worldwide. Early identification of COVID-19 is very crucial, as it keeps the infected people isolated from other people, thus minimizing the risk of further transmission. The standard diagnostic approach is based on RT-PCR. However, due to the scarcity of PCR kits in some regions and the costs associated with this technique, there is a growing demand for alternative solutions. Recently, diagnosis of COVID-19 by medical imaging has been recognized as a valid clinical practice. Meanwhile, the massive increase in COVID-19 cases has put considerable pressure on radiologists responsible for interpreting these scans. This paper introduces an automated detection approach as a rapid alternative for COVID-19 diagnosis. We present a deep CNN model to differentiate between normal and pneumonia cases, as well as patients with COVID-19. Our approach is based on EfficientNet-B7 architecture and improved with Squeeze and Excitation block as an attention mechanism. In addition, we propose an innovative architecture that combines CNN with SVM to achieve the best performance. Experimental results show that the proposed framework provides better performance than existing SOTA methods, with an average accuracy of 97.50%, while the precision and recall of COVID-19 are both 100%.

Keywords: COVID-19, Chest X-ray, CNN, EfficientNet-B7, Squeeze and Excitation block, Support Vector Machine

1. Introduction

Coronaviruses form a diverse group of respiratory infectious viruses causing a variety of disorders, from the common cold to severe pneumonia. In 2019, a new coronavirus variant appeared in Wuhan, China, identified as SARS-CoV-2. The appearance of this variant, commonly known as COVID-19, has prompted a global state of emergency, straining healthcare systems worldwide [1]. This perilous viral infection has had a major impact on many people's lives, being five times more fatal than influenza, with serious illnesses and deaths [2]. Similar to many viruses, COVID-19 can mutate when it propagates among individuals, giving rise to variants like Delta and Omicron. These variants can present variable levels of transmission, symptoms, and disease severity. COVID-19 generally incubates from 2 to 14 days after initial exposure. and can infect either the superior respiratory system (nose, throat,..etc) or the inferior respiratory system (windpipe and lungs) [3].

Currently, there is no antiviral that has been clinically approved for COVID-19, underlining the crucial role of rapid diagnosis and isolation of infected individuals in limiting the spread of the virus. The most reliable diagnostic technique is the RT-PCR test [4]. Nevertheless, this test still has limitations in terms of availability and high costs. Recently, several clinical studies have shown that chest imaging can also help to identify and diagnose COVID-19 at an early stage [5]. CT scans have emerged as the first option for lung imaging diagnosis of suspected COVID-19 cases, particularly when patients present with respiratory disorders. However, CT scanning is not widely used due to the practical problems of transporting patients to CT centers and the associated risk of radiation exposure, particularly for pregnant women and children [6]. On the other hand, chest X-rays (CXR) are a promising alternative that reduces the risk of radiation exposure and offers a faster, and more effective alternative.

Recently, the rapid propagation of the virus has imposed a significant workload on radiologists, increasing the potential risk of medical diagnosis errors. For this reason, computer-assisted diagnosis systems can be an interesting alternative, providing radiologists with a valuable tool for faster, and better interpretation of chest radiographs. A variety of automated techniques for medical imaging analysis have achieved significant progress in recent years, but deep learning (DL) was the major innovation. Its application in medicine has led to many research studies focusing on the identification of various illnesses, including brain disorders from MRI [7], various brain disorders from EEG [8], and skin diseases [9]. Moreover, in certain cases of medical imaging, the literature has shown that the classification performance of DL models can achieve or surpass the performance of clinicians [10]. In the framework of chest

E-mail address: rami.naidji@univ-sba.dz

radiography, deep learning has shown its ability to diagnose various thoracic pathologies, such as pneumothorax, pneumonia, effusion [11], and recently COVID-19 [12].

Motivated by these findings, we conducted a study for the classification of chest radiographic images to distinguish normal cases from pneumonia or COVID-19 using the largest COVID-19 open-source dataset. We present a deep model based on EfficientNet-B7 architecture and attention mechanism. We chose EfficientNet-B7 as it is known to be highly accurate, faster, and smaller than the best CNN model (66M parameters vs Gpipe with 556M) [13]. Meanwhile, the attention mechanism enables the network to highlight only significant features of the chest X-ray, while reducing the importance of features that are not significant to the classification task. Furthermore, to achieve the best performance, we train an SVM algorithm with deep features extracted from our CNN architecture. The relevance and novelty of this study lie in the fact that it integrates a DL model with a traditional ML algorithm. The results demonstrate that our solution outperforms state-of-the-art (SOTA) approaches.

The main novel contribution of this work can be illustrated in the following way:

1)- The adaptation of EfficientNet-B7 architecture for COVID-19 diagnosis, namely : COVID-EfficientNetB7-Basic.

2)- The use of an attention mechanism to enhance the deep extracted features. We add the SE block at the top of EffecientNetB7.

3)- The proposal of a combination of CNN and SVM.

4)- A comparative study with SOTA techniques.

The paper is structured as follows. Section 2 summarizes existing studies. Section 3 describes our methodology. Section 4 presents the evaluation of our approaches using the COVIDx dataset, as well as the experimental results. Then, in section 5 we conclude and suggest potential improvements.

2. Related work

The spread of coronavirus worldwide has made chest imaging a crucial diagnostic technique that can control the propagation of the disease. Due to the increasing need for rapid radiological analysis, the artificial intelligence research community has focused its efforts on developing AI algorithms to improve diagnosis precision [14].

Most studies on the identification of COVID-19 from X-rays have been inspired by earlier research on the detection of pneumonia. Researchers have used a variety of deep learning frameworks, including CNN, RNN, autoencoder, as well as hybrid networks such as CNN-RNNs and CNN-AEs [15]. Farooq et al [16] presented a pre-trained ResNet-50 [17] on ImageNet [18] (called COVIDResNet). The head of the model is substituted by an alternative head made up of a series of adaptive pooling, dropout, batch normalization, and fully connected layers. This fine-tuning produced remarkable results, with an accuracy of 96.23% for all classes, including normal, bacterial, viral, and COVID-19 cases, with only 41 epochs on the COVIDx dataset. Haghanifar et al, [19] presented a hierarchical DL model. In the early stage, the model differentiates chest radiographs into two categories: normal or radiographs with pneumonia. Then, in the next stage, it improves its classification by distinguishing between COVID-positive cases and pneumonia cases. This approach yielded impressive results, with a global accuracy of 87.21% and f1-scores of 92.00% for COVID-19 and 85.00% for pneumonia. Rahimzadeh et al [20] trained multiple deep-learning models to classify chest radiographs into distinct classes: normal, pneumonia, and COVID-19. They developed a novel approach by fusing Xception [21] and ResNet50V2 [22] models. Their approach obtained the better performance, with the highest accuracy observed on a test set of 11,302 images. They yielded an impressive overall accuracy of up to 91.40% across all classes, along with an accuracy of 99.50% specifically for COVID-19 cases. Apostolopoulos et al. [23], in their research, evaluated the performance of SOTA pre-trained CNN architectures introduced in medical imaging classification in recent years. These architectures included VGG19 [24], MobileNetV2 [25], Inception [26], Xception [21], Inception-ResNet v2 [27]. Fine-tuning performance has shown the effectiveness of using CNNs with transfer learning method for the classification task. Their best-performing model achieved remarkable accuracy, and recall rate of 96.78%, and 98.66%, respectively. Turkoglu [28] proposed COVIDetectioNet. In this approach, features were used from AlexNet. Then, a feature selection was carried out using the Relief algorithm to identify the most relevant features. Finally, the selected features were used for classification using SVM. Remarkably, the author achieved an outstanding accuracy rate of 99.18% on a dataset containing 609 images, including 21 cases of COVID-19 pneumonia. Pan et al [29] designed multi channel feature DNN. To address the issue of imbalanced data, they employed an oversampling technique to mitigate the deviation in prior probability of MFDNN algorithm. As a result, they achieved an accuracy of 93.19%.

While the approaches mentioned have shown promising results, it is essential to note that many of these research studies performed their evaluation on relatively small datasets, often using a restricted set of COVID-19 images. In addition, given the recent appearance of the pandemic, standardized research databases did not yet exist. As a consequence, each study relied on a different database, which contributed to a lack of uniformity between research studies. In contrast, in [12], the authors used a comprehensive dataset comprising 13,800 X-ray images from 13,645 individuals, with 182 images of COVID-19 pa-

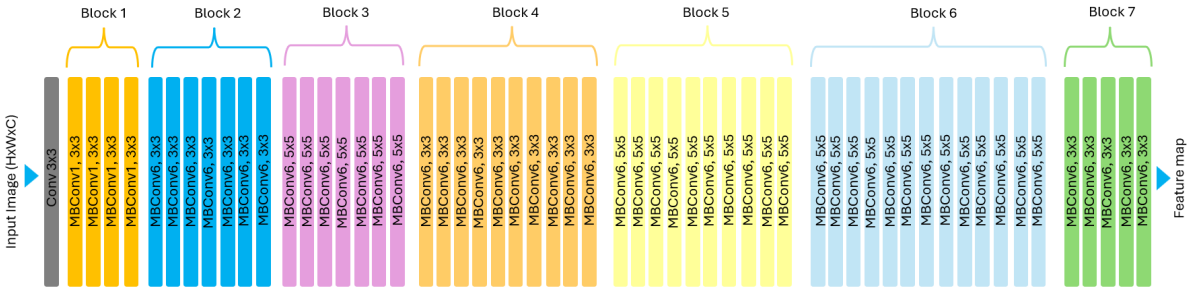

Figure 1. The EffecientNet-B7 architecture.

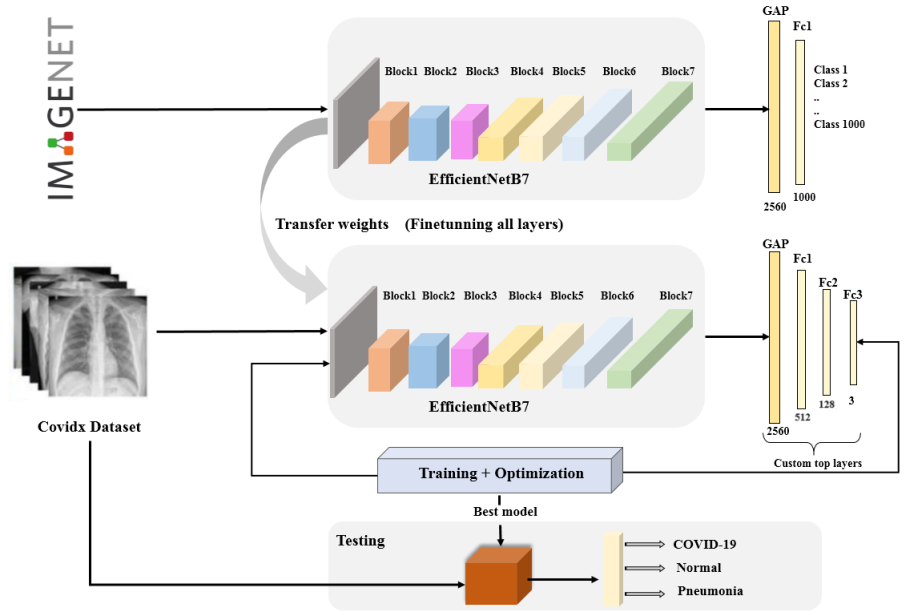

Figure 2. The overview of the proposed architecture COVID-EffecientNetB7-Basic (Baseline). EffecientNetB7 is used as a features extractor, followed by fully connected layers that perform the classification task.

tients. In addition, they designed a deep CNN model called COVID-Net to distinguish normal cases from pneumonia and COVID-19. In particular, COVID-Net presents significantly reduced architectural and computational complexity in comparison to well-known architectures such as VGG-19 [24] and ResNet-50 [17], while providing better results. They achieved a test accuracy of 93.30% and a COVID-19 recall rate of 91.00%.

With the appearance of this new dataset, Khobahi et al [30] introduced a novel semi-supervised learning approach employing AutoEncoders, and a CNN network as a classifier. The authors reported an average accuracy of 93.50%. Ratul et al [31] presented a novel CNN model based on MobileNet enhanced with residual connections and attention mechanism. They reached an average accuracy of 95.30%. Chowdhury et al [32] presented ECOVNET, an ensemble of CNN based on EfficientNet architectures (from B0 to B5) [13]. They obtained an average accuracy of 97.00%, while the precision and recall of COVID-19 are both 100%. Also, Luz et al [33] introduced a series of CNNs built upon the EfficientNet architecture. They proposed a hierarchical classification strategy employing two classifiers: one focused on distinguishing between Normal and Pneumonia classes, and another dedicated to discriminating between different pneumonia types, including COVID-19 and non-COVID pneumonia. Their approach achieved an average accuracy of 93.50%. Bhadouria et al [34] have introduced a framework based on the following steps: first, images are pre-processed using histogram equalization and segmented with a U-network. Next, features are extracted via VGG-16 and an SVM model is used for classification. They obtained an average accuracy of 96.36%. Szczepanski et al [35] used multi task neural network based on U-Net and ResNet50. They reached an average accuracy of 90.30%. Rahhal et al [36] adopted a siamese deep learning architecture based on Vision Transformer. They achieved an average accuracy of 94.62%. Rangel et al [37] presented a Capusl CNN

network. They used dilation rate instead of max-poling as a novel contribution. Their approach achieved an average accuracy of 90.00 %. Ullah et al [38] presented a dense CNN network based on attention mechanism. The authors reported an average accuracy of 97.22%, with 96.87% the COVID-19 detection rate. Anwar et al [39] introduced a new self-supervised learning approach with a vision transformer model. Their approach reached an average accuracy of 96.00%.

3. The proposed solution

This section, describes our method for COVID-19 diagnosis. We present the approach and implementation details.

A. *Adapted Efficient-B7 for COVID-19 diagnosis*

Recently, Tan et al [13] explored the relationship between the depth and width of CNN models, revealing a more efficient strategy for designing models with reduced parameters while improving classification accuracy. The authors presented a pioneering paper describing a novel class of CNN known as EfficientNet models.

The fundamental component of the EfficientNet model family [13] is the Mobile Inverted Bottleneck Convolution (MBConv), enhanced through SE optimization. EfficientNet variants integrate a variable set of these blocks, derived from concepts initially introduced in MobileNet [25]. In addition, EfficientNet introduces Swish as a new activation function. It takes a comparable form to ReLU and LeakyReLU, sharing some of their performance advantages. Furthermore, the key importance of EfficienNet lies in the fact that it uses a new scaling method to generate seven distinct models from EfficientNet-B0 (baseline), namely: EfficientNet-B1, EfficientNet-B2,..., EfficientNet-B7. As we move from EfficientNet-B0 to EfficientNet-B7, several key attributes, including model width, depth, resolution, and complexity, progressively increase, resulting in improved accuracy. These architectures have shown superior accuracy to a large number of convolutional neural networks while maintaining considerably enhanced computational capabilities.

In our study, we use the EfficientNet-B7 as a backbone in our architecture, Figure 1. The features generated are then transmitted through our custom top layers, which consist of three fully connected layers. The initial layer, denoted as ”fc1” with 512 units, and the subsequent layer labeled ”fc2” with 128 units, use ReLU, while the last layer, ”fc3,” is activated with softmax, which plays an important role in the ultimate stages of the classification process (Figure 2). The proposed architecture has been trained using the transfer learning technique. This approach aims to take advantage of knowledge transfer between domains [40]. It allows certain fundamental features such as shapes, corners, edges, and intensity to be shared among various tasks, which speeds up the training process and improves overall model performance.

Instead of initializing the network weights randomly, the model weights are initialized with those derived from a pre-trained model on ImageNet [18]. ImageNet's pre-trained weights have shown robust performance in image analysis and demonstrated superior results in several studies involving chest X-ray data, as reported in [10] and [41]. These weights only provide an initial starting point and are fine-tuned to suit the specific characteristics of the X-ray images.

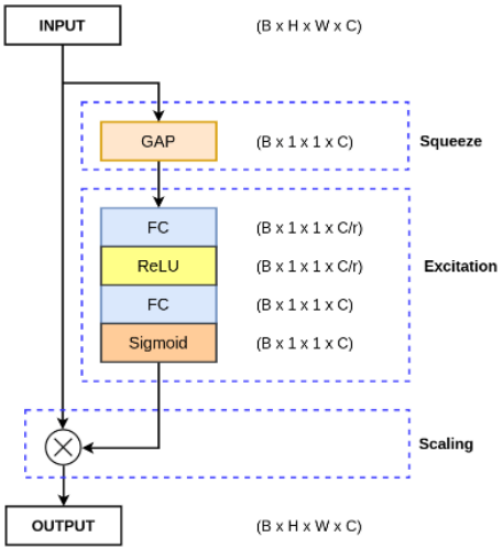

Figure 3. The Squeeze and Excitation block. Input: features map. B: batch size, H: height, W: width, C: number of channels; r: reduction factor. Output: re-weighted features map.

B. *Attention mechanism*

The convolution process generates a feature map with variable channel numbers, determined by the filter parameters. Typically, these channels are processed uniformly in the following operations, implying an equal level of importance for each of them. However, this uniform processing approach may not be the most optimal. As a significant improvement, the attention mechanism can be used to introduce a capability into the neural network, enabling it to prioritize particular features by giving them distinct weights [42]. This attention technique has found applications in various areas of deep learning, including image recognition [43], speech recognition [44], object detection [45], and NLP [46].

To enhance the performance of our proposed framework, we incorporated the SE block [47] as a channel attention mechanism with the feature extractor; EfficientNet-B7. Figure 3 shows the global structure of this block. It's a component designed to enhance the representational power of convolutional neural networks (CNNs) by explicitly modeling channel-wise relationships and recalibrating feature maps. It consists of two main operations: squeezing, which aggregates global information across spatial dimensions, and exciting, which learns channel-wise dependencies

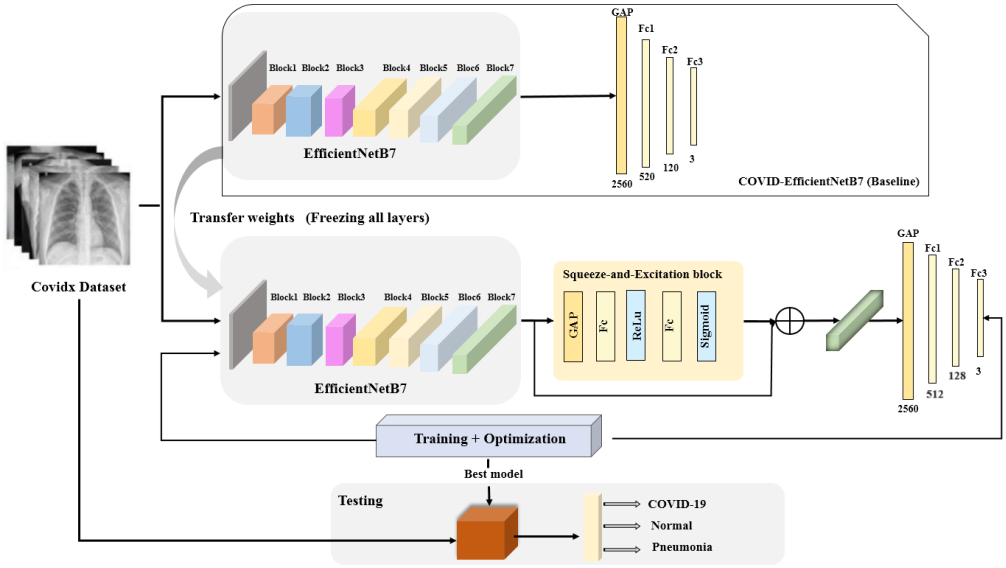

Figure 4. The overview of the proposed architecture COVID-EffecientNetB7-Attn. The SE block is used to improve the network's representational capability by focusing on relevant features and removing less useful ones.

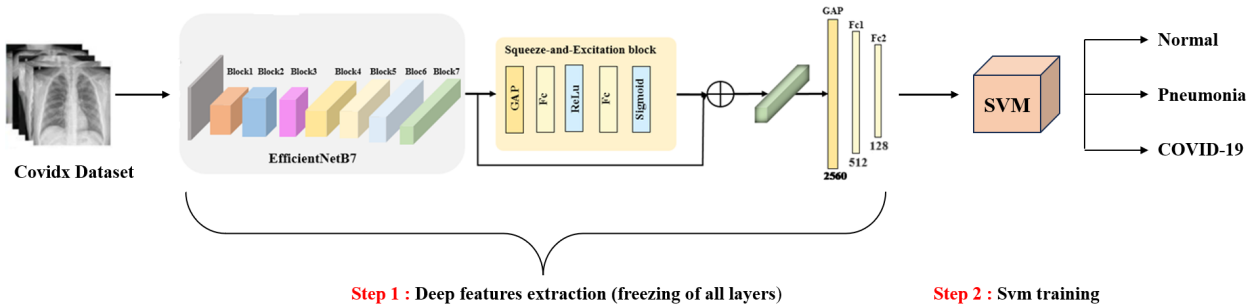

Figure 5. The overview of the proposed architecture COVID-EffecientNetB7-Attn based SVM.

to generate importance scores. This lightweight attention mechanism improves feature discrimination without significantly increasing model complexity, leading to better performance in image classification, object detection, and semantic segmentation tasks. Instead of dealing with all channels in the same way, the SE block allows the network to assign distinct weights to each channel, to reinforce the interdependencies between these channels. The advantage of this approach is that these weight values are learned just like any other parameter during the learning process. Finally, the SE block yields an output shape of (1 x 1 x channels), representing the weights associated with each channel, which are used to rescale the initial features. Following this operation, the feature map retains only relevant information, improving the representation capability of the network.

Given a CXR image, feature map is extracted from our EffecientNetB7 model which was previously trained in a supervised way on the COVIDx dataset. Then, this feature map underwent a "squeeze" operation, through a global average pooling layer, yielding a channel descriptor. This descriptor captures a representative embedding of the overall distribution of feature responses across the channel dimension enabling all network layers to access information about the global receptive field. Following aggregation, we introduce an "excitation" operation, comprising two fully connected layers. The first layer significantly reduces the number of features by a factor "r" (where choosing r = 8 strikes a good balance in terms of accuracy and complexity). This layer uses a ReLU activation function, while the second layer employs a sigmoid function to transform the values into a range between 0 and 1. The excitation operation essentially take the simple form of a self-gating control mechanism. It received the embedding as input and generated a set of modulation weights, with one weight per channel. These weights are then scaled to the feature maps derived from the EfficientNetB7 model. If a weight value is close to 0, it implies that the corresponding channel is less critical, leading to a reduction in the feature channel value. Conversely, if the weight value approaches 1, it indicates that the channel is significant, and pixel values are preserved

to a greater extent. After the full process, we transmit the scaled feature map to the custom top layers through three fully connected layers, which perform classification task, Figure 4.

C. Combined COVID-EffecientNetB7-Attn with SVM model

To improve the efficiency of our proposed architecture, we introduce an advanced hybrid architecture that merges a SOTA CNN model with a traditional classifier; SVM [48]. Building on the capabilities of these two proven image recognition components, our hybrid system exploits the unique strengths of each classifier, resulting in an enhanced architecture. We use our architecture EfficientNet-B7-based SE block as a feature extractor. Then, the deep features extracted from the penultimate layer (128 D) are used as input to train the SVM classifier, Figure 5. In other words, we replace the softmax function with a multiclass linear SVM. The key idea behind the application of SVM is to use a supervised learning algorithm to identify the hyperplane that allows optimal separation of features in a high-dimensional space [49]. Along this learning process, we aim to minimize a margin-based loss, in contrast to the more traditional approach of using a cross-entropy loss.

The motivation for using linear SVM in our study stems from its simplicity and effectiveness in classification tasks. Unlike ANNs, which involve complex architectures and parameter settings, SVM offers a straightforward approach by finding the optimal hyperplane to separate classes in our data. In addition, SVM tends to generalize well and is less susceptible to over-fitting, especially in high-dimensional data scenarios, which is a common problem with ANNs. Consequently, given the trade-off between complexity and performance, linear SVM appears to be a practical choice for our classification challenge. To determine appropriate values for the penalty parameter C, we performed a small grid search over the values 0.0001, 0.001, 0.01, 0.1, 10, and 100 ultimately selecting c=10. Then, the proposed hybrid architecture is trained using the validation set.

4. Experiments

The experiments carried out in this study are based on the COVIDx dataset [12]. We conduct model testing using the test set made available by the authors, which includes 400 images of three different groups: COVID-19, pneumonia, and normal lung. Then, we performed extensive cross-validation on the entire dataset; 30530 images.

A. Dataset

We used the COVIDx [12] dataset, the largest X-ray open-source dataset for automatic COVID-19-19 identification. The authors have used various publicly available data repositories to build up the data set: RSNA Pneumonia Detection Challenge dataset [50], Covid-chest x-ray-dataset [51], Actualmed [52], Figure 1 [53], BIMCV-COVID19+ [54], COVID-19 Radiography Database V3 [55], [56].

The experiments conducted in this study rely on the latest subdivision provided by the authors, known as COVIDx V9A, Table I. This version includes three distinct categories: COVID-19, pneumonia, and normal. To our knowledge, it's the largest open-source dataset for automatic COVID-19 identification. Sample images from this benchmark dataset are shown in Figure 6. The training set includes 30,130 images from 16,424 patients, distributed as follows: 8,085 images representing healthy human lungs, 5,555 images representing lungs with pneumonia, and 16,490 images representing lungs with COVID-19. On the other hand, the test set consists of 400 images derived from 378 patients including 100 images of normal lungs, 100 images of pneumonia cases, and 200 images of COVID-19.
It is essential to note that the COVIDx dataset is evolving continuously [12]. New patient cases are constantly being incorporated and regularly made available to the public. This study is based on a specific snapshot of the dataset, representing its status at a given time.

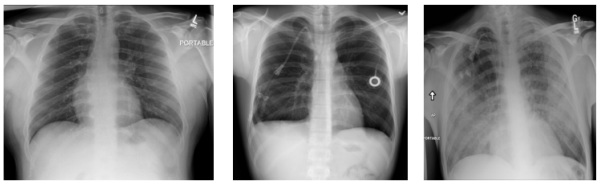

Figure 6. Example of chest radiograph samples from the COVIDx dataset.

TABLE I. Distribution of COVIDx V9A images between the classes and partitions of the training and test sets [12]. Number of patients between brackets.

Subset	COVID-19	Normal	Pneumonia	Total
Train	16490 (2808)	8085 (8085)	5555 (5531)	30130 (16424)
Test	200 (178)	100 (100)	100 (100)	400 (378)

B. Environment setup

The experiments were conducted in a virtual environment with the following specifications:

- RAM of 12.70 GB.
- TPU with 8 cores.

C. Implementation details

In the model training process, we randomly split the training set in a train set and a validation set. Specifically, we kept 20% of data for the validation process, as shown in the table II. We made sure there was no overlap of patients between the sets. The primary reason for this division is to assess the model's generalization on unseen data, simulating real-world scenarios. The training set is used to tune the model's parameters, while the validation set provides an independent dataset to evaluate its performance and tune hyperparameters. In addition, to optimize these hyper-parameters, we carry out experiments involving the adjustment of learning rate (lr), dropout rate, and epochs. We also explored different activation functions, including ReLU, Swish, and others, eventually choosing ReLU as the

TABLE II. Distribution of images between the classes and their partitions within each set.

Subset	COVID-19	Normal	Pneumonia	Total
Train	13182	6479	4443	24104
Validation	3308	1606	1112	6026
Test	200	100	100	400

activation function. In the pre-processing step, we resize each image to 600x600, which corresponds to the resolution parameter of the EfficientNet-B7. Next, we normalize the image vectors by dividing all pixel values by 255.

Due to the nature of our multi-class challenge, we use categorical cross-entropy. It is specifically designed for scenarios where the target variable has multiple classes, and it computes the dissimilarity between the predicted probabilities and ground truth distributions, as described below:

$$\text{LOSS} = -\sum_{i=1}^{\text{output size}} y_i \cdot \log \hat{y}_i \tag{1}$$

Where $\hat{y}_i$ is the predicted value, and y_i is the relevant class label. The dimension is the number of classes, specifically 3 in this scenario. By optimizing the model to minimize this loss, we encourage it to learn accurate and well-calibrated probability distributions over the classes.

For both architectures, COVID-EffecientNetB7-Basic and COVID-EffecientNetB7-Att, we train the entire model with Adam optimizer for only 30 epochs on TPU using a batch size of 64. Each batch of input data is distributed among the TPU replicas, with each replica receiving an input size of 8. To enhance the model's generalization, we introduce a 1% dropout on the fc1 and fc2 layers, and we initialize the lr to 0.0001 for the baseline architecture and 0.001 for COVID-EfficientNetB7-Attn. Then, reduce the lr by a coefficient of 10 when the validation loss reaches a plateau after every 5 epochs. The selected model is the one associated with the least validation loss. Table III illustrates the Hyper-parameters configuration of our models.

TABLE III. Hyper-parameters configuration.

	COVID-EffecientNetB7 Basic	C0VID-EffecientNetB7 Attn
Optimizer	Adam	Adam
Epoch	30	30
Batch size	8 * 8 TPU cores	8 * 8 TPU cores
Learning rate	0.0001	0.001
Dropout	1% after dense layers	

The loss curves for both architectures are presented in Figures 7a and 8a. The training loss is a metric that quantifies how well the model is learning from the training data data during the training phase. On the other hand, the validation loss is crucial for assessing how well it generalizes to new, unseen data. The gap between these curves serves as an indicator of potential overfitting. Notably, both curves exhibit a gradual decline, signifying a continuous improvement in our models as they undergo training. A favorable fit is characterized by minimal disparity between the final training and validation loss values, indicative of effective generalization. In Figures 7b and 8b as well, we can observe the training and validation accuracy. For each architecture, model performance steadily advances with training experience but eventually reaches a plateau, suggesting that further learning becomes challenging.

D. Results

To assess the effectiveness of the suggested approaches, we used a set of measures: accuracy, precision, recall (sensitivity), and F1 score. This approach enables us not only to evaluate our architectures but also to facilitate comparisons with other relevant studies. These metrics are described as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$F1 - Score = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

$$Avg_{Accuracy} = \frac{\sum^{n} \frac{TP+TN}{TP+TN+FP+FN}}{n} \tag{6}$$

Accuracy is a measure of the succeeded predictions compared to the total number of predictions. Precision quantifies the proportion of successfully predicted true positives among all predicted positive instances. Recall represents the rate at which the model correctly recognizes positive samples. The F1 score tries to strike a balance between precision and recall. Avg Accuracy is a measure of the performance of the classification method.

1) COVID-EffecientNetB7-Basic (without attention)

We evaluate our baseline architecture on the test set given by the authors of the dataset [12]. The performance are summarized in Table IV. As shown, the accuracy achieves 100%, 78.00%, and 94.00% for COVID-19, Normal, and Pneumonia categories, without data augmentation and with an imbalanced training set. The proposed CNN shows a high recall rate for COVID-19; 100%, which is particularly relevant, as it indicates the efficiency of our solution to minimize the number of missed cases. In addition, we obtain a good precision rate of 97%, indicating a minimal number of false-positive detections of COVID-19 cases.

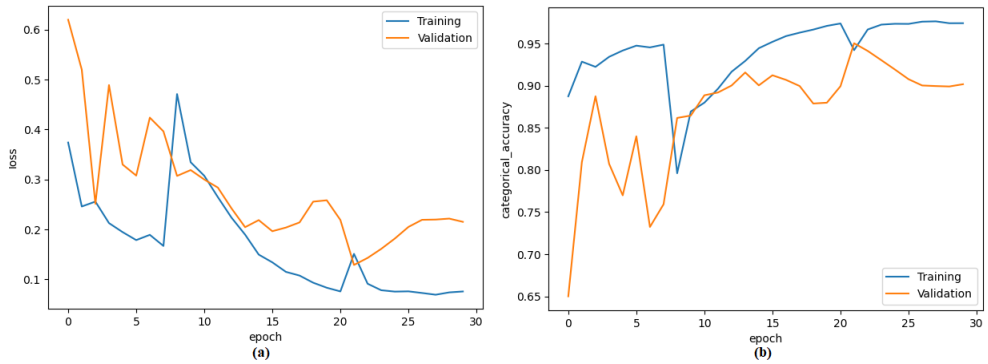

Figure 7. Learning curves of COVID-EffecientNetB7-Basic.(a) Loss. (b) Accuracy.

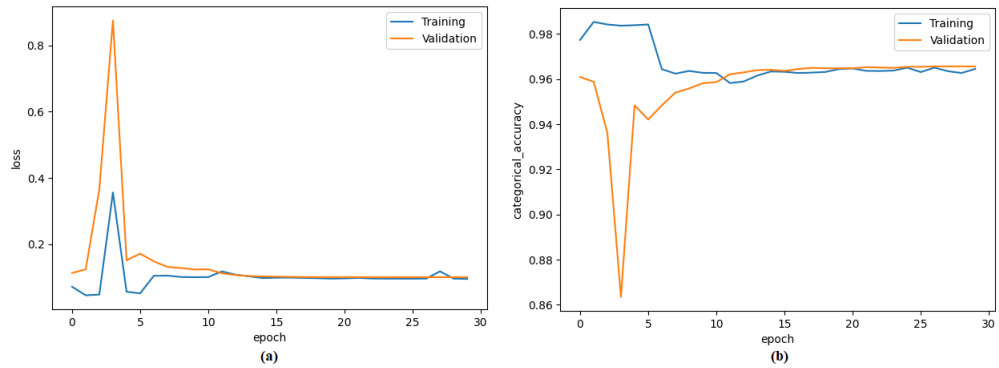

Figure 8. Learning curves of COVID-EffecientNetB7-Attn. (a) Loss. (b) Accuracy.

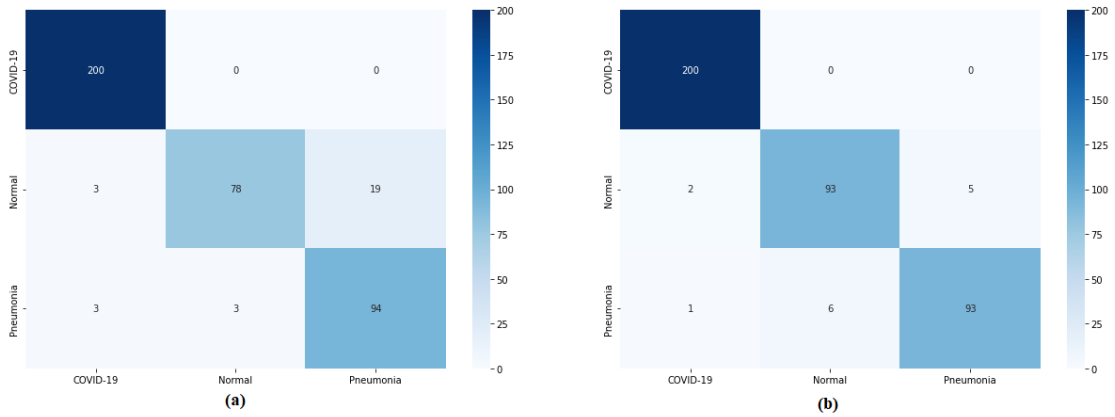

Figure 9. Confusion matrix, (a): COVID-EffecientNetB7-Basic (baseline) architecture, (b): COVID-EffecientNetB7-Attn.

TABLE IV. Classification results of the proposed architectures on COVIDx dataset.

Architecture	Class	Precision (%)	Recall (%)	F1-score (%)	$Avg_{Accuracy}$ (%)
COVID-EffecientNetB7-Basic (without attention)	COVID-19	97.00	**100**	99.00	93.00
	Normal	**96.00**	78.00	86.00	
	Pneumonia	83.00	**94.00**	88.00	
COVID-EffecientNetB7-Attn (with attention)	COVID-19	**99.00**	**100**	**99.00**	**96.50**
	Normal	94.00	**93.00**	**93.00**	
	Pneumonia	**95.00**	93.00	**94.00**	

TABLE V. Patients and images distribution in each fold.

Fold	COVID-19		Normal		Pneumonia		Overall(images)
	Patients	Images	Patient	Images	Patients	Images	
Fold 1	293	1663	810	810	576	580	3053
Fold 2	297	1670	820	820	563	563	3053
Fold 3	296	1663	820	820	564	570	3053
Fold 4	301	1672	800	800	579	581	3053
Fold 5	303	1673	827	827	550	553	3053
Fold 6	300	1667	835	835	545	551	3053
Fold 7	302	1675	818	818	560	560	3053
Fold 8	298	1668	807	807	576	578	3053
Fold 9	295	1666	850	850	536	537	3053
Fold 10	301	1673	798	798	582	582	3053

TABLE VI. Performance metrics for each class label of the COVID-EffecientNetB7-Basic with 10-fold cross-validation strategy.

Fold	COVID-19			Normal			Pneumonia		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Fold 1	99.00	100.00	99.00	95.00	97.00	96.00	95.00	90.00	92.00
Fold 2	100	100	100	95.00	94.00	95.00	92.00	93.00	92.00
Fold 3	99.00	99.00	99.00	94.00	93.00	94.00	89.00	91.00	90.00
Fold 4	100	100	100	95.00	92.00	94.00	89.00	93.00	91.00
Fold 5	100	99.00	99.00	88.00	94.00	91.00	90.00	83.00	86.00
Fold 6	99.00	100.00	99.00	93.00	97.00	95.00	96.00	86.00	91.00
Fold 7	100	99.00	99.00	94.00	94.00	94.00	91.00	93.00	92.00
Fold 8	99.00	100	99.00	92.00	89.00	91.00	87.00	88.00	88.00
Fold 9	99.00	100	100	93.00	95.00	94.00	92.00	88.00	90.00
Fold 10	100	100	100	92.00	97.00	94.00	95.00	87.00	91.00

TABLE VII. The support-weighted average performance of the COVID-EffecientNetB7-Basic within cross-validation.

	Fold1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Average ± SD
Accuracy	97.00	97.00	96.00	97.00	94.00	97.00	96.00	95.00	96.00	97.00	96.20 ± 0.97
Precision	97.00	97.00	96.00	97.00	95.00	97.00	97.00	95.00	96.00	97.00	96.40 ± 0.80
Recall	97.00	97.00	96.00	97.00	94.00	97.00	96.00	95.00	96.00	97.00	96.20 ± 0.97
F1-score	97.00	97.00	96.00	97.00	94.00	97.00	97.00	95.00	96.00	97.00	96.30 ± 1.00

In cases where the data set is relatively small, as is the case in our study (with only 400 images in the test set), the use of cross-validation becomes crucial to ensure a fair evaluation of the classification method. To evaluate the effectiveness of the solution and ensure robustness in its performance assessment, we employ a 10-fold cross-validation approach as our chosen validation method. We first combined the train and test sets, giving a total of 30,530 images. Then, we divided this dataset into ten equal subsets, each fold containing 3,053 images, as shown in Table V. Importantly, we ensured that there was no patient overlap between these subsets. In each iteration, we keep only 10% of data for the test. Specifically, the 10th subset served as the test set, while the other nine subsets constituted collectively

the training data. A summary of the performance metrics for each class at each fold is presented in the table VI. While the support-weighted average of performance measures is mentioned in Table VII. An average accuracy of 96.20%, precision of 96.40%, sensitivity of 96.20%, and F1-score of 96.30% were obtained.

2) Enhanced Efficient-B7 with attention mechanism

To assess the contribution and impact of the SE block, we perform an ablation study. The initial model is called COVID-EfficientNetB7-Basic, while the second model is called COVID-EfficientNetB7-Attn. The classification results on the test set are presented in table IV. Compared to the proposed baseline, the use of SE block enables a significant improvement in prediction performance, with an accuracy of 96.50%, while the precision and sensitivity for COVID-19 identification are 100% and 99.00% respectively. As illustrated by the confusion matrix in Fig. 9, all 200 COVID cases are accurately identified. Similarly, for both normal patients and those with pneumonia, 93 out of 100 cases are correctly classified, with only 7 cases being misclassified into other categories. Experimental results demonstrate that COVID-EfficientNetB7-Attn offers superior performance for all metrics. The introduction of the SE block results in a significant improvement, increasing recall by around 15% for the normal class, precision by around 2% for the COVID-19 class, and by 12% for the pneumonia class, underlining its effectiveness. More specifically, precision and recall rates for the COVID-19, normal, and pneumonia classes can be summarized as follows P: (99.00/94.00/95.00) and R: (100/93.00/93.00) respectively. Furthermore, the benefits of incorporating an attention mechanism into the EfficientNetB7 architecture are evident in the significant improvement in F1 scores across all classes compared to using EfficientNet-B7 alone. This improvement exceeded the baseline approach by at least 7% for the normal class and 6% for the pneumonia class.

The SE block measures the relevance of each feature channel through a learning process. By selectively amplifying informative features and suppressing less relevant ones, the SE block facilitates the extraction of more discriminative representations from the input data. This enables the model to better capture complex patterns and variations, ultimately leading to improved accuracy. Also, through global average pooling followed by excitation, the SE block integrates global context information into the feature representation process. By considering the relationships between different channels, the model can capture higher-level semantic information that may be crucial for accurate prediction. In contrast, without an attention mechanism, the model treats all features equally, potentially diluting the importance of crucial information and hindering its ability to discern subtle patterns.

To demonstrate the effectiveness of our approach, Figure 10 provides a visualization of the attention weights learned

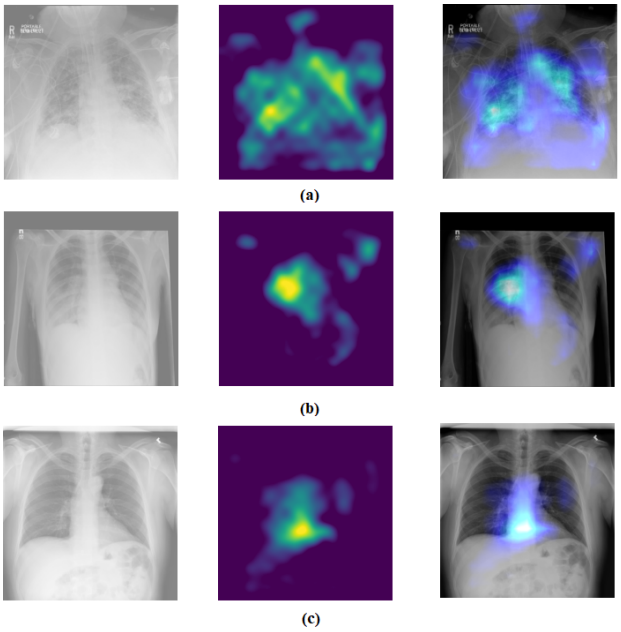

Figure 10. Activation attention maps for various CXR images. The heatmap highlights the important areas detected by the model. (a) : COVID-19 sample. (b) : Pneumonia sample. (c): No findings (normal).

by the architecture, where intense colors show higher attention levels. In Figure 10a, which represents a COVID-19 case, our model accurately localizes bilateral shadow patterns in the lungs (known as Ground Glass Opacity), a significant sign of COVID-19 [57]. The presence of blue color in the lungs indicates abnormalities in the chest X-ray image. Figure 10b depicts a pneumonia case, where the heatmap distinctly highlights the white spots (opacity) in the lungs. As well, Figure 10c shows normal cases, where the heatmap predominantly focuses on regions outside the lungs or in proximity to the heart. These results demonstrate the importance of the attention mechanism in extracting significant features and enhancing the performance of the classification task.

3) Combined COVID-EffecientNetB7-Attn with SVM model

The performances of the hybrid architecture are summarized in Figure 11 and Table IX. Using the penultimate fully connected layer of the COVID-EffecientNetB7-Attn architecture with a linear SVM classifier produces significantly improved results compared to the conventional softmax function. The hybrid model achieves the best performance as it combines the key characteristics of CNN, SE Block, and SVM. We can see that COVID-EffecientNetB7-Attn-based SVM outperforms the other methods on several metrics. We achieve an accuracy of 97.50%, while the precision and recall rate of COVID-19 detection both reach 100%. In addition, a considerable improvement in F1-score is achieved in all classes. The lower recall rate (93.00%) on the pneumonia class is probably a consequence of the limited number of cases in our patient population.

TABLE VIII. A comparison of the proposed architectures with state-of-the-art on COVIDx dataset. The year of the research study is in brackets.

Method	Number of images		COVID-19		$Avg_{Accuracy}$ (%)
	COVID	Non-COVID	Precision (%)	Recall (%)	
COVID-Net. (2020) [12]	358	13617	98.90	91.00	93.30
CoroNet. (2020) [30]	99	18430	90.00	90.00	93.50
RAM-Net. (2020) [31]	358	13617	99.00	92.00	95.30
ECOVNet. (2020) [32]	589	14904	**100**	**100**	97.00
Luz et al. (2021) [33]	183	13587	**100**	96.80	93.90
Bhadouria et al. (2021) [34]	470	2000	98.02	98.71	96.36
POTHER. (2022) [35]	353	14357	100	95.00	90.30
ViT model. (2022) [36]	358	13604	92.84	90.00	94.62
DR CapsNet. (2023) [37]	1770	14919	-	-	90.00
DAM-Net. (2023) [38]	2442	14920	96.87	95.54	97.20
ViT-S + SS-CXR. (2023) [39]	16194	14299	99.70	95.50	96.00
COVID-Efficient-Attn (Ours)	**16690**	13840	99.00	**100**	96.50
COVID-Efficient-Attn + Linear SVM (Ours)	**16690**	13840	**100**	**100**	**97.50**

TABLE IX. Performance of the COVID-EffecientNetB7-Attn based SVM on COVIDx dataset.

Class	Precision (%)	Recall (%)	F1-Score (%)	$Avg_{Accuracy}$ (%)
COVID-19	100	100	100	
Normal	93.00	97.00	95.00	97.50
Pneumonia	97.00	93.00	95.00	

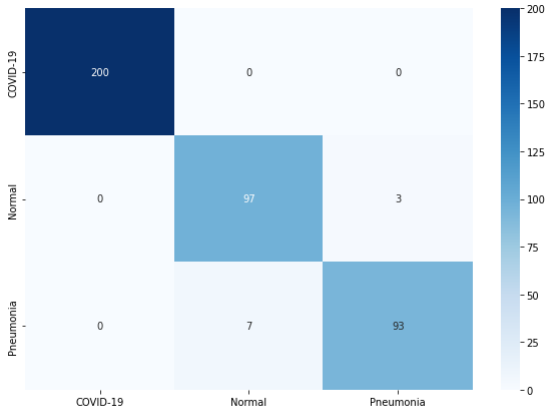

Figure 11. Confusion matrix of the proposed COVID-EffecientNetB7-Attn based SVM.

4) Comparative study

In this research, we have conducted a comparative study of the results obtained by the proposed architectures against SOTA methods. Table VIII presents an overview of the comparison across several metrics; average accuracy, sensitivity, and precision. The results range from an accuracy of 90.00% in [37] to 97.20% in [38]. The proposed architecture COVID-Efficient-Attn-based linear SVM shows better results than all SOTA methods. It achieves a relatively high average accuracy of 97.50%, precision, and sensitivity of 100% for COVID-19 detection. This underlines the importance of feature extraction using an optimized model that takes into account three critical aspects: increased depth and width, as well as wider image resolution. These considerations enable finer details to be captured, ultimately improving classification accuracy. On the other hand, it is essential to highlight the challenge posed by the limited availability of COVID-19 images in the majority of these studies. Many SOTA methods rely on small datasets containing a limited number of images, with only a few COVID-19-positive patient samples. In contrast, our proposed architecture has been developed and evaluated using a large COVID-19 benchmark, which contains the largest number of COVID-19-positive cases (16690 X-ray images). It is therefore better adapted to new and unseen images than any existing work in the literature.

5. Conclusion and futur works

Through this study, we introduce an artificial intelligence-based approach that improves the radiological diagnosis of COVID-19. We presented a robust deep-learning architecture designed to accurately identify visual features of COVID-19 and influenza virus pneumonia in radiological images. Our solution holds promise for complementing and aiding radiological diagnosis by facilitating accurate and timely differentiation between COVID-19 and influenza virus pneumonia. Furthermore, the proposed approach is flexible, making it applicable to a wide range of classification tasks beyond this specific application.

In the future, potential directions of research could involve incorporating lung segmentation techniques to further improve the model's performance. In addition, when more data become available, it will be possible to extend our experimental work by approving the method on larger datasets. Furthermore, we look forward to extending our diagnostic system to other lung pathologies, like chronic obstructive pulmonary disease, bronchitis, etc, to assess its validity and reliability in a broader clinical context.

Data availability : The dataset generated during and/or analyzed during the current study is available to download from the following links: https://www.kaggle.com/datasets/andyczhao/covidx-cxr2 and https://github.com/lindawangg/COVID-Net.

Declaration

Conflict of interest : The authors declare that they have no conflicts of interest to report regarding this study.

References

[1] world health organization, "Novel coronavirus(2019-ncov) situation report - 11," 01 2020. [Online]. Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200131-sitrep-11-ncov.pdf

[2] L. Piroth, J. Cottenet, A.-S. Mariet, P. Bonniaud, M. Blot, P. Tubert-Bitter, and C. Quantin, "Comparison of the characteristics, morbidity, and mortality of COVID-19 and seasonal influenza: a nationwide, population-based retrospective cohort study," *The Lancet Respiratory Medicine*, vol. 9, no. 3, pp. 251–259, Mar. 2021. [Online]. Available: https://doi.org/10.1016/s2213-2600(20)30527-0

[3] M. S. Razai, K. Doerholt, S. Ladhani, and P. Oakeshott, "Coronavirus disease 2019 (covid-19): a guide for uk gps," *BMJ*, vol. 368, 2020.

[4] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, and R. Agha, "World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19)," *International Journal of Surgery*, vol. 76, pp. 71–76, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1743919120301977

[5] M.-Y. Ng, E. Lee, J. Yang, F. Yang, X. Li, H. Wang, M. Lui, C. Lo, B. S. T. Leung, P. Khong, C. Hui, K.-y. Yuen, and M. Kuo, "Imaging profile of the COVID-19 infection: Radiologic findings and literature review," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 1, p. e200034, Feb. 2020. [Online]. Available: https://doi.org/10.1148/ryct.2020200034

[6] D. HE, W. CG, and G. FV, "The risks of radiation exposure related to diagnostic imaging and how to minimise them," *BMJ*, vol. 342, 2011.

[7] N. Ghassemi, A. Shoeibi, and M. Rouhani, "Deep neural network with generative adversarial networks pre-training for brain tumor classification based on mr images," *Biomedical Signal Processing and Control*, vol. 57, pp. 101 678–101 688, 11 2019.

[8] H. He, X. Liu, and Y. Hao, "A progressive deep wavelet cascade classification model for epilepsy detection," *Artificial Intelligence in Medicine*, vol. 118, p. 102117, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S093336572100110X

[9] N. Hameed, A. M. Shabut, and M. A. Hossain, "Multi-class skin diseases classification using deep convolutional neural network and support vector machine," in *2018 12th International Conference on Software, Knowledge, Information Management Applications (SKIMA)*, 2018, pp. 1–7.

[10] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Y. Ding, A. Bagul, C. P. Langlotz, K. S. Shpanskaya, M. P. Lungren, and A. Y. Ng, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," *CoRR*, vol. abs/1711.05225, 2017. [Online]. Available: http://arxiv.org/abs/1711.05225

[11] 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," *CoRR*, vol. abs/1705.02315, 2017. [Online]. Available: http://arxiv.org/abs/1705.02315

[12] 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," *Scientific Reports*, vol. 10, no. 1, p. 19549, Nov 2020.

[13] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 09–15 Jun 2019, pp. 6105–6114.

[14] H. Asad, H. Azmi, P. Xi, A. Ebadi, S. Tremblay, and A. Wong, "Covid-19 detection from chest x-ray images using deep convolutional neural networks with weights imprinting approach," *Journal of Computational Vision and Imaging Systems*, vol. 6, pp. 1–3, 01 2021.

[15] A. Shoeibi, M. Khodatars, M. Jafari, N. Ghassemi, D. Sadeghi, P. Moridian, A. Khadem, R. Alizadehsani, S. Hussain, A. Zare, Z. A. Sani, F. Khozeimeh, S. Nahavandi, U. R. Acharya, and J. M. Gorriz, "Automated detection and forecasting of covid-19 using deep learning techniques: A review," *Neurocomputing*, vol. 577, p. 127317, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231224000882

[16] M. S. Farooq and A. Hafeez, "Covid-resnet: A deep learning framework for screening of covid19 from radiographs," *ArXiv*, vol. abs/2003.14395, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:214727906

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[18] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[19] A. Haghanifar, M. M. Majdabadi, Y. Choi, S. Deivalakshmi, and S. Ko, "COVID-CXNet: Detecting COVID-19 in frontal chest x-ray images using deep learning," *Multimedia Tools and Applications*, Apr. 2022. [Online]. Available: https://doi.org/10.1007/s11042-022-12156-z

[20] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2," *Informatics in Medicine Unlocked*, vol. 19, p. 100360, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352914820302537

[21] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, jul 2017, pp. 1800–1807.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 630–645.

[23] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, Apr. 2020. [Online]. Available: https://doi.org/10.1007/s13246-020-00865-4

[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.

[26] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

[27] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 02 2017, p. 4278–4284.

[28] M. Turkoglu, "COVIDetectioNet: COVID-19 diagnosis system based on x-ray images using features selected from pre-learned deep features ensemble," *Applied Intelligence*, vol. 51, no. 3, pp. 1213–1226, Sep. 2020. [Online]. Available: https://doi.org/10.1007/s10489-020-01888-w

[29] L. Pan, B. Ji, H. Wang, L. Wang, M. Liu, M. Chongcheawchamnan, and S. Peng, "MFDNN: multi-channel feature deep neural network algorithm to identify COVID19 chest x-ray images," *Health Information Science and Systems*, vol. 10, no. 1, Apr. 2022. [Online]. Available: https://doi.org/10.1007/s13755-022-00174-y

[30] 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*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:216055366

[31] M. A. Rab Ratul, M. Tavakol Elahi, K. Yuan, and W. Lee, "Ramnet: A residual attention mobilenet to detect covid-19 cases from chest x-ray images," in *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2020, pp. 195–200.

[32] N. K. Chowdhury, M. A. Kabir, M. M. Rahman, and N. Rezoana, "ECOVNet: a highly effective ensemble based deep learning model for detecting COVID-19," *PeerJ Computer Science*, vol. 7, p. e551, May 2021. [Online]. Available: https://doi.org/10.7717/peerj-cs.551

[33] E. Luz, P. Silva, R. Silva, L. Silva, J. Guimarães, G. Miozzo, G. Moreira, and D. Menotti, "Towards an effective and efficient deep learning model for COVID-19 patterns detection in x-ray images," *Research on Biomedical Engineering*, Apr. 2021. [Online]. Available: https://doi.org/10.1007/s42600-021-00151-6

[34] H. S. Bhadouria, K. Kumar, A. Swaraj, K. Verma, A. Kaur, S. H. Sharma, G. Singh, A. Kumar, and L. M. de Sales, "Classification of covid-19 on chest x-ray images using deep learning model with histogram equalization and lungs segmentation," *ArXiv*, vol. abs/2112.02478, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:244909381

[35] T. Szczepański, A. Sitek, T. Trzciński, and S. Płotka, "Pother: Patch-voted deep learning-based chest x-ray bias analysis for covid-19 detection," in *Computational Science – ICCS 2022*, D. Groen, C. de Mulatier, M. Paszynski, V. V. Krzhizhanovskaya, J. J. Dongarra, and P. M. A. Sloot, Eds. Cham: Springer International Publishing, 2022, pp. 441–454.

[36] M. M. Al Rahhal, Y. Bazi, R. M. Jomaa, A. AlShibli, N. Alajlan, M. L. Mekhalfi, and F. Melgani, "Covid-19 detection in ct/x-ray imagery using vision transformers," *Journal of Personalized Medicine*, vol. 12, no. 2, 2022. [Online]. Available: https://www.mdpi.com/2075-4426/12/2/310

[37] G. Rangel, J. C. Cuevas-Tello, M. Rivera, and O. Renteria, "A deep learning model based on capsule networks for covid diagnostics through x-ray images," *Diagnostics*, vol. 13, no. 17, 2023. [Online]. Available: https://www.mdpi.com/2075-4418/13/17/2858

[38] Z. Ullah, M. Usman, S. Latif, and J. Gwak, "Densely attention mechanism based network for covid-19 detection in chest x-rays," *Scientific Reports*, vol. 13, no. 1, Jan. 2023. [Online]. Available: http://dx.doi.org/10.1038/s41598-022-27266-9

[39] S. M. Anwar, A. Parida, S. Atito, M. Awais, G. Nino, J. Kitler, and M. G. Linguraru, "Spcxr: Self-supervised pretraining using chest x-rays towards a domain specific foundation model," *arXiv 2211.12944*, 2023. [Online]. Available: https://arxiv.org/abs/2211.12944

[40] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[41] E. Hussain, M. Hasan, M. A. Rahman, I. Lee, T. Tamanna, and M. Z. Parvez, "Corodet: A deep learning based classification for covid-19 detection using chest x-ray images," *Chaos, Solitons & Fractals*, vol. 142, p. 110495, 2021.

[42] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, "Classification of remote sensing images using efficientnet-b3 cnn model with attention," *IEEE Access*, vol. 9, pp. 14 078–14 094, 2021.

[43] M. Pedersoli, T. Lucas, C. Schmid, and J. Verbeek, "Areas of attention for image captioning," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017.

[44] P.-W. Hsiao and C.-P. Chen, "Effective attention mechanism in dynamic models for speech emotion recognition," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018.

[45] Y. Zhu, C. Zhao, H. Guo, J. Wang, X. Zhao, and H. Lu, "Attention CoupleNet: Fully convolutional attention coupling network for object detection," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 113–126, Jan. 2019. [Online]. Available: https://doi.org/10.1109/tip.2018.2865280

[46] A. Galassi, M. Lippi, and P. Torroni, "Attention in natural language processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 10, pp. 4291–4308, Oct. 2021. [Online]. Available: https://doi.org/10.1109/tnnls.2020.3019893

[47] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.

[48] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: https://doi.org/10.1007/bf00994018

[49] H. Basly, W. Ouarda, F. E. Sayadi, B. Ouni, and A. M. Alimi, "CNN-SVM learning approach based human activity recognition," in *Lecture Notes in Computer Science*, 2020, pp. 271–281.

[50] E. Tsai, S. Simpson, M. P. Lungren, M. Hershman, L. Roshkovan, E. Colak, B. J. Erickson, G. Shih, A. Stein, J. Kalpathy-Cramer, J. Shen, M. A. Hafez, S. John, P. Rajiah, B. P. Pogatchnik, J. T. Mongan, E. Altinmakas, E. Ranschaert, F. C. Kitamura, L. Topff, L. Moy, J. P. Kanne, and C. Wu, "Medical imaging data resource center (midrc) - rsna international covid radiology database (ricord) release 1c - chest x-ray, covid+ (midrc-ricord-1c)," 2021.

[51] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future," *arXiv 2006.11988*, 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset

[52] A. G. Chung, "Actualmed covid-19 chest x-ray data initiative," 2020. [Online]. Available: https://github.com/agchung/Actualmed-COVID-chestxray-dataset

[53] A. G. Chung., "Figure1 covid-19 chest x-ray data initiative," 2020. [Online]. Available: https://github.com/agchung/Figure1-COVID-chestxray-dataset

[54] M. de la Iglesia Vayá, J. M. Saborit-Torres, J. A. Montell Serrano, E. Oliver-Garcia, A. Pertusa, A. Bustos, M. Cazorla, J. Galant, X. Barber, D. Orozco-Beltrán, F. García-García, M. Caparrós, G. González, and J. M. Salinas, "Bimcv covid-19+: a large annotated dataset of rx and ct images from covid-19 patients," 2021. [Online]. Available: https://dx.doi.org/10.21227/m4j2-ap59

[55] M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. A. Emadi, M. B. I. Reaz, and M. T. Islam, "Can ai help in screening viral and covid-19 pneumonia ?" *IEEE Access*, vol. 8, pp. 132 665–132 676, 2020.

[56] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. A. Maadeed, S. M. Zughaier, M. S. Khan, and M. E. Chowdhury, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest x-ray images," *Computers in Biology and Medicine*, vol. 132, p. 104319, May 2021. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2021.104319

[57] H. Y. F. Wong, H. Y. S. Lam, A. H.-T. Fong, S. T. Leung, T. W.-Y. Chin, C. S. Y. Lo, M. M.-S. Lui, J. C. Y. Lee, K. W.-H. Chiu, T. W.-H. Chung, E. Y. P. Lee, E. Y. F. Wan, I. F. N. Hung, T. P. W. Lam, M. D. Kuo, and M.-Y. Ng, "Frequency and distribution of chest radiographic findings in patients positive for covid-19," *Radiology*, vol. 296, no. 2, pp. E72–E78, 2020, pMID: 32216717.

Mohamed Rami Naidji is currently a Ph.D. student in the field of artificial intelligence. He specializes in the development and application of machine learning and deep learning techniques for the analysis of clinical data. His academic journey has been driven by an unwavering commitment to the intersection of cutting-edge technology and healthcare.

Pr. Zakaria Elberrichi , is a professor and researcher with over 30 years of experience in the fields of artificial intelligence (AI), machine learning, and deep learning. Currently serving as a faculty member at Djillali Liabes University, Sidi belabbes, Algeria. He has made significant contributions th these rapidly evolving domains through his pioneering research and dedicated teaching. His work has been published extensively, with over 50 papers in prestigious journals and conference proceedings. In addition, Pr. Zakaria Elberrichi has supervised and mentored numerous doctoral students, guiding them through their research projects and fostering the next generation of AI experts. Under his supervision, more than 15 Ph.D. students have completed their doctoral studies.