A Hierarchical Classification System for the Detection of Covid-19 from Chest X-Ray Images

Meghna P Ayyar Jenny Benois-Pineau Akka Zemmari
Univ. Bordeaux, CNRS, LaBRI, UMR 5800, F-33400 Talence, France
meghna-parameswaran.ayyar@etu.u-bordeaux.fr
{jenny.benois-pineau,akka.zemmari}@u-bordeaux.fr

Abstract

With the ever-increasing cases of the Covid-19 pandemic, it is important to leverage deep learning methods to create tools that can aid in relieving the pressure that is put on the limited resources in most developing countries. In this work, we propose a hierarchical classification system for the classification of Covid-19 from Chest X-Ray (CXR) images following a recent proposal of massive use of this modality instead of CT. The system composed of multiple binary classifiers outperforms a tailor-made multi-class classifier COVID-Net. We also show that using well-known established deep learning frameworks combined with a global attention mechanism outperforms the baseline COVID-Net specifically designed for the classification of Covid-19 from CXR images. Our method shows approximately a 4% improvement in the sensitivity to Covid-19 detection from 91% of COVID-Net to 96%. Using popular networks with the possibility of cross-domain transfer learning ensures that the designing and training times are reduced. Furthermore, well-established frameworks can be faster adapted into an application in clinical practice.

1. Introduction

The novel COVID-19 or SARS-Cov-2 is an infectious disease that has been declared as a pandemic by WHO in March 2020 [36]. First reported in Wuhan, China at the end of 2019 [12], it has had devastating effects on human life and the economy. Many nations are still combating its proliferation and are facing a shortage of resources. The crucial step to control and stop Covid-19 is to detect infected patients effectively and impose immediate isolation.

Currently, Reverse Transcription Polymerase Chain Reaction (RT-PCR) or gene sequencing for respiratory or blood specimens are the standard screening methods for COVID-19 [30]. However, these methods have a long turnaround time and are not sufficient when there are exponential increases in daily cases. The effectiveness of chest radiography imaging (X-Ray, Computed Tomography (CT) imaging) for the detection of infection[34, 15, 23] makes it a suitable alternative in the screening process and for the design of adapted therapy.

A Chest X-Ray (CXR) image affected by COVID-19, shows patchy or diffuse reticular-nodular opacities and consolidation, with basal, peripheral and bilateral predominance[8]. The major abnormalities observed in CT images are ground-glass opacity, consolidation and interlobular septal thickening in both lungs[32]. Fig.1 shows a sample CXR image that has been annotated by an expert radiologist to show the primary regions for diagnosis making. Though CT is a more precise modality for decision making by medical experts[25], CXR images are useful, specifically in low-income countries as they are cheaper and more widely available, and XR examinations of patients can be done with a portable equipment. Furthermore, patients with acute respiratory syndrome are difficult to manipulate to be placed into CT scanners and is easier with XR equipment. Finally, CXR images enable rapid triaging as they can be completed much faster than a CT scan. With a high volume of patients, this is essential to relive pressure on the available resources.

In CXR, Covid-19 syndrome and atypical pneumonia appear similar and experts cannot always make distinction just by observing the CXR image[25]. Hence the the automatic classification on images with Deep Neural Networks (DNN) is necessary not only for massive screening, but also for highlighting specific patterns on this modality.

Recent medical studies [21, 22, 29] amongst others have attempted to develop specific Deep Neural architectures for the automatic detection of COVID-19 on CXR images and some open datasets have become available for research purposes [5, 17, 7]. The main contribution of our work are as follows:

- We define a set of hierarchical rules to create a multi-level classification pipeline for the three classes -

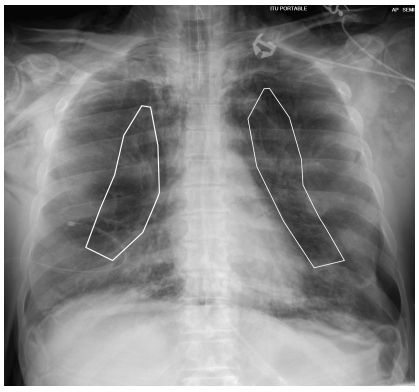


Figure 1. Sample annotation of Chest X-Ray (CXR) image that has been affected with the regions showing presence of pneumonia that could be Covid-19 as annotated by a radiologist

Normal (Healthy), Pneumonia and COVID-19. The pipeline is composed of binary classifiers which proved to be more efficient than a multi-class classifier on medical images [1].

- We propose using well-known networks with robust open-source components and complementing them with attention mechanisms instead of tailor-made networks for Covid-19 detection. We believe that such an approach facilitates the deployment of AI tools in clinical practice.
- We present detailed results of all the binary classifiers that have been tested and compared them with the baseline result of the tailor-made network COVID-Net [29].

The rest of the paper is organized as follows. In section 2 we present recent studies on the classification of CXR images in COVID-19 studies. Our hierarchical classification pipeline with an attention mechanism is presented in section 3. In section 4 results are presented and discussed. Finally, section 5 concludes the work and presents its perspectives.

2. Related Work

Deep learning approaches have been widely used to identify different thoracic diseases, including pneumonia [33, 28, 31] due to the availability of large-scale annotated benchmarking datasets like CheXpert [13] and ChestXray14 [31] to name a few. To design a specific architecture tailored for Covid-19 detection, a large dataset is required. Unlike other diseases, for a relatively new disease like Covid-19, the availability of sufficiently large datasets is scarce e.g. COVID-19 Image Data Collection [7] and ActualMed COVID-19 Chest X-Ray Dataset Initiative[5].

Many works like [20, 14, 22, 3] have tested various transfer learning strategies to improve the effectiveness of the approach in classifying the CXR images as either healthy, pneumonia or COVID-19 cases. The different approaches that have been adopted are bag of visual words[16], enhanced feature selection methods like Cuckoo Search Optimization algorithms (CS) [35] and self-organization maps (SOM) [19] to name a few.

Recently, the COVIDx Chest X-Ray dataset [29] was made available as an open-access benchmark dataset, which contains one of the largest numbers of publicly available COVID-19 positive cases under the COVID-Net Open Source Initiative [footnote cite]. COVID-Net [29] was also proposed as the benchmark network on COVIDx achieving 91% sensitivity for the Covid-19 detection. In addition to highly imbalanced classes of Normal(healthy), Pneumonia and Covid-19, another challenge with the datasets are that generally with CXR images atypical pneumonia is very similar to Covid-19 making these two classes not easy to be separated by classifiers. Next, the COVID-Net is a tailor-made Deep CNN for multi-class classification on CXR. There is no guarantee that this architecture is optimal for the large amount of newly coming data. Hence our claim is that one can build an efficient classification framework from already available architectures that have been well-studied in large scale tasks combined with popular attention mechanisms. Hence in the following, we propose our classification framework and compare it on COVID-Net on the same benchmark - COVIDX dataset.

3. Hierarchical Covid-19 classification pipeline

The proposed pipeline consists of multiple binary classifiers trained to distinguish between the different classes present in the dataset. It combines a form of ensemble methods of majority voting and a set of defined hierarchical rules to create a multi-level classifier instead of a single multi-class classifier. For imbalanced datasets, a low-populated class might not be well modelled by a multi-class classifier, which is typically the Covid-19 class in this case. Consequently, a multi-class hierarchical pipeline was designed as shown in Fig 2. The pipeline consists of four trained binary classifiers:

- Classifier 1: Normal v/s Positive. The Normal is the healthy lung images and Positive are the images with either Pneumonia or COVID-19.
- Classifier 2: Normal v/s COVID-19
- Classifier 3: Normal v/s Pneumonia
- Classifier 4: Pneumonia v/s COVID-19

Instead of learning to distinguish three classes, the binary classifiers are fine-tuned to distinguish between two classes at a time. The following rules have been used to make the final prediction for a given image:

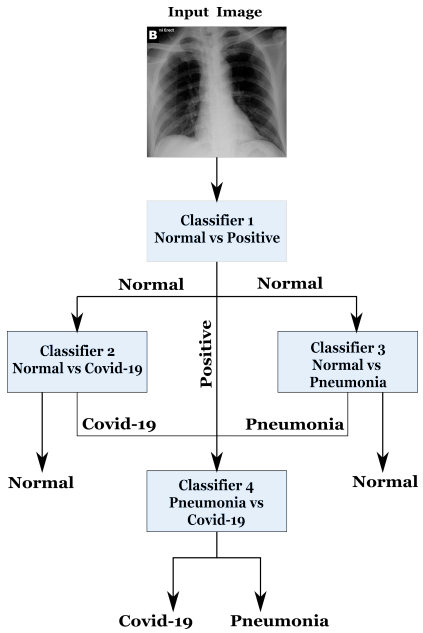


Figure 2. The proposed Multi-level classification pipeline of binary classifiers for- Covid-19, Pneumonia and Normal classes

- If Classifier 1 predicts a normal class- then Classifiers 2 and 3 are tested. If the majority of these three classifiers predicts normal then the result is given as normal.
- If Classifier 1 predicts a positive class or there is no majority vote as Normal from the classifiers 1, 2 and 3 then, Classifier 4 is tested. Its answer is either Pneumonia or Covid-19. This decision is finally checked by the majority vote between classifiers 2,3 and 4.

The impact of an undetected Covid-19 is much higher due to the adverse effects of infection progression and the risk of transmission. In a similar classification problem of Alzheimer's disease from sMRI images [2] observe that majority voting works better than other schemes and hence we have chosen it to be a part of our rules. The first rule is focused on ensuring that false-negative error i.e. the prediction of a case as normal is greatly reduced. The second rule is aimed at learning the distinction between the Pneumonia and Covid-19 classes.

3.1. CNN architectures used for COVID-19 classification

Nowadays there exists a variety of well established Deep Neural Network frameworks that are used for classification purposes. These networks have been tested across different domains and on large-scale datasets[10, 26, 24].

[21] observe that pre-trained networks that have been trained on ImageNet [9] show good performance when adapted for CXR images. This strategy is called **Cross-domain transfer learning**, as the target domain, X-Ray images strongly differ from the source domain (ImageNet) comprising general-purpose colour images. In the previous work [1], it was shown that intra-domain transfer, when realized between different modalities of the medical images on the same subjects, outperforms the cross-domain knowledge transfer. Unfortunately, different imaging modalities are generally not available for Covid-19 screening. Therefore, cross-domain transfer learning with pre-trained networks on a very large (ImageNet) database as stated in [21] is used in the actual work.

A tailor-made network requires a large dataset to train and the architecture needs to be validated thoroughly before it can be adopted for a real-world application. Using well-known network architectures makes their adoption easier and can be trusted more as they have been widely assessed. We aim to have easily fine-tuned and better-performing networks to ensure that the training process is faster than creating a network from scratch. ResNet50 [10] and InceptionV3 have been chosen [26] over other popular networks. The choice of the networks is based on their better performance on the ImageNet dataset. These networks are architecturally different to check which works better on CXR images. They have lesser amount of parameters to train and need lower computational power when compared to their deeper variants or the VGG networks [24].

Further, CNN classifiers are designed by fixing several hyperparameters. A part of these parameters constitutes the number of filters for each convolutional layer. It is extremely difficult to perform a full search in this parameter subspace to select an optimal number of filters to train for each convolutional layer. Another way of network optimization consists in the weighting of feature channels resulting from convolutions thus giving different importance to each filter. This mechanism is called "global attention" as the importance of a channel is assigned globally to all features in it. We use Squeeze-Excitation (SE) [11] based attention to make the pre-trained networks more robust to the dataset. Furthermore, in the classification task with such an importance as Covid diagnosis, the information support for medical decision has to be endorsed on only well-studied approaches. SE mechanism is a well studied mechanism with available software to add to any CNN for channel- wise weighting.

3.1.1 Global Attention based on Squeeze-Excitation(SE) Block

In a CNN the tensor that is computed at a 2D-convolutional (conv) layer is of a dimension of $(H \times W \times C)$ where $H \times W$ is the spatial dimension of the feature map at the layer and C is the number of channels each of them resulting from a convolution of the input with one of the filters of the layer. The *Squeeze-Excitation (SE)* [11] based attention module can be

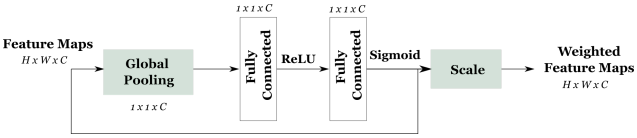


Figure 3. Squeeze-Excitation (SE) Block, $H \times W$ - the height and width of the feature map and C - channels at the layer

used in tandem with any convolutional layer to weight each channel in the conv layer to remove redundancy. Channels with a low contribution to the output will have lower attention. The overall framework for the SE block is summarized by Fig 3.

The SE-attention block consists of the following

- **Squeeze Module**: It squeezes the information of each of the feature maps from the previous convolution layer to a singular value per channel using Global Average Pooling (GAP) Eq 1 where $a_c(i, j)$ is the activation of that particular position (i, j) in a channel feature map. Here i, j range over the dimensions H and W of the map respectively. Input feature maps of size $(H \times W \times C)$ are reduced to a tensor of $(1 \times 1 \times C)$.

$$z_c = \frac{1}{H \times W} \sum_{H}^{i=1} \sum_{W}^{j=1} a_c(i, j) \tag{1}$$

- **Excitation Module**: It learns the adaptive scaling weights for the channels during training. The excitation module is a Multi-Layer Perceptron (MLP) comprised of a single hidden layer with a reduction factor r. Thus, the input layer to the MLP has C neurons, the hidden layer with $\frac{C}{r}$ neurons and the output with C neurons. The output layer has a sigmoid (Eq 2) activation in order to scale the weights between 0 to 1 to get in the tensor of size $(1 \times 1 \times C)$.

$$sig(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

- **Scale Module**: The scaling weights are applied to the feature maps obtained from the convolutional layer by an element-wise multiplication scaling each $(H \times W)$ map with the corresponding scaling weight.

The SE based attention has a very low computation burden and would not affect the training time. It is a simple but effective attention mechanism and Fig 4 & 5 shows the scheme of the placement of the SE blocks which were added for the ResNet50 and InceptionV3 networks. In our case, the SE block is added only with the last residual and inception blocks.

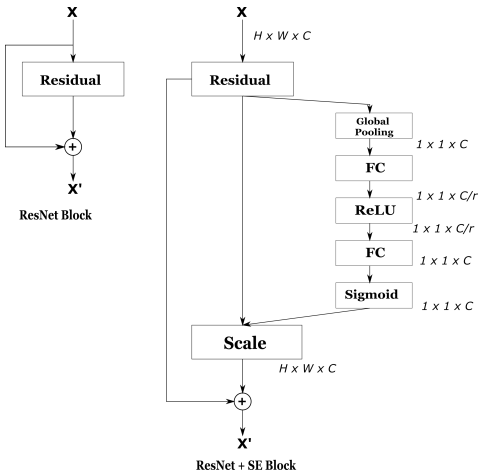


Figure 4. Schema of the original ResNet residual block (left), the residual block after adding the SE Block (right)

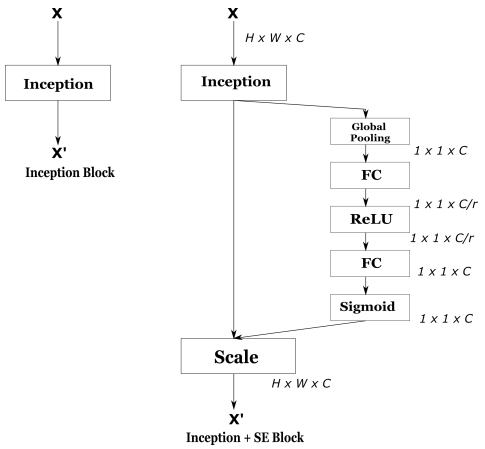


Figure 5. Schema of the original Inception block (left), the inception block after adding the SE Block (right)

4. Results & Discussion

In this section we will briefly present the COVIDx dataset, explain the training of our hierarchical classifiers and present the results in detail.

4.1. Dataset Details

We have used the COVIDx dataset[29] that consists of 13,975 Chest X-Ray (CXR) images across 13,870 patient cases. The dataset has been proposed as an open-access benchmark dataset that was created by combining five other publicly available data repositories namely, a) COVID-19 Image Data Collection [7], b) Figure 1 COVID-19 Chest X-ray Dataset Initiative [6], c) ActualMed COVID-19 Chest X-ray Dataset Initiative, established in collaboration with ActualMed[5], d) RSNA Pneumonia Detection Challenge dataset, which used publicly available CXR data from [18], and e) COVID-19 radiography database [17]. Fig 6 illustrates different images taken from each of these datasets to

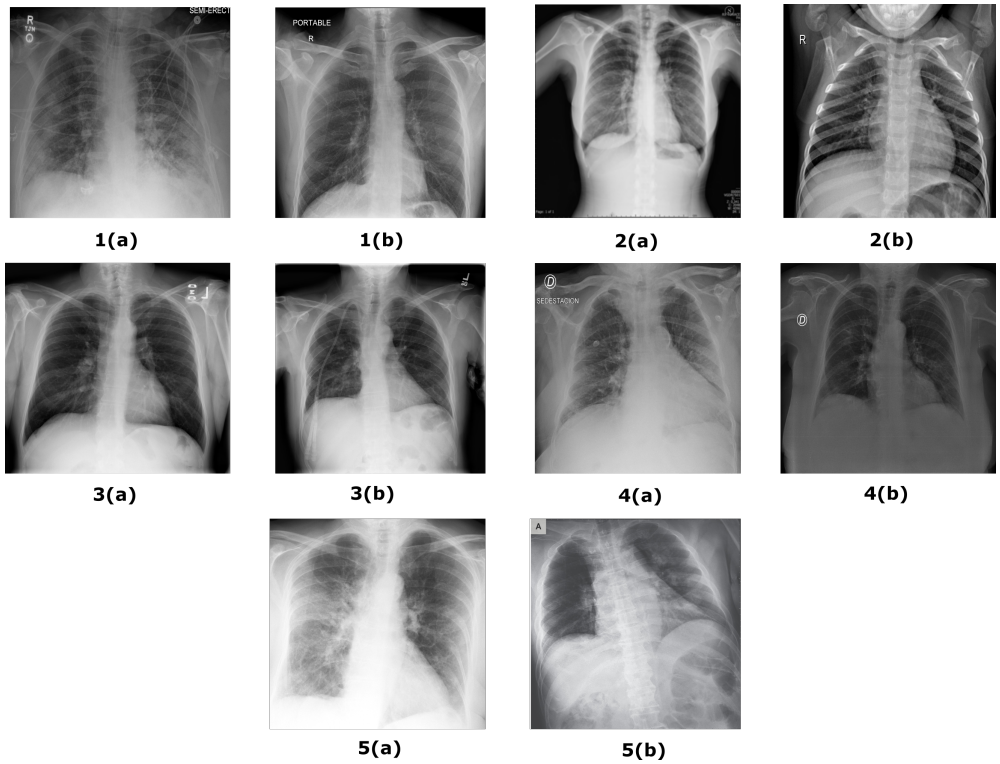


Figure 6. Sample images from COVIDx dataset from multiple repositories, 1. COVID-19 Image Data Collection [7], 2. COVID-19 Chest X-Ray Dataset Initiative[6], 3. RSNA Pneumonia Detection challenge dataset[18], 4. ActualMed COVID-19 Chest X-Ray Dataset Initiative[5], 5. COVID-19 radiography database[17]

create the final COVIDx dataset.

The COVIDx dataset contains three classes: Normal (X-Rays that do not contain Pneumonia or Covid-19), Pneumonia (X-Rays that have some form of bacterial or viral pneumonia, but no Covid-19) and Covid-19 (X-Rays that are Covid-19 positive). The method for the creation of the official dataset has been summarized in [1]. The COVIDx dataset has two versions a. Full COVIDx b. Official COVIDx that differ only in the test set. The official COVIDx is used by the authors to report the performance of their network. During the creation of the dataset, some images of the COVID-19 radiography database [17] were unavailable. Consequently, the final distribution of the dataset varies from the official figures and is shown in Fig 7. In our case, the training set has approximately 1.11% lesser samples of Covid-19 images than the official COVIDx dataset. But, the test set of official COVIDx and our case are the same. Further, a 90-10% validation split has been made for network training.

4.2. Training Procedure

In this section we briefly present the data pre-processing and the parameters that have been used for network training.

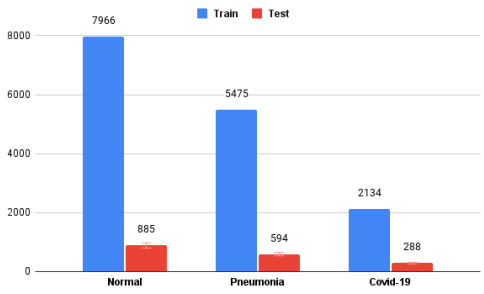
[1] https://github.com/lindawangg/COVID-Net/blob/master/docs/COVIDx.md

4.2.1 Data pre-processing

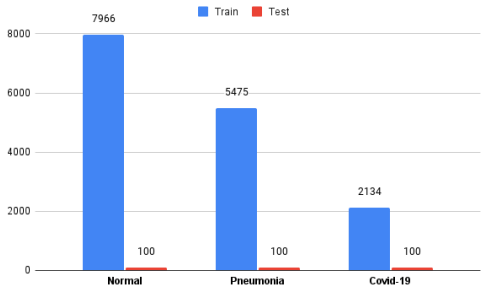
The COVIDx dataset consists of 2D images in the png and jpg formats. As the dataset is a collection of images from multiple sources images do not have the same resolution. Therefore we scaled them to fit the input resolution of the network under consideration. Firstly a uniform scaling of each image accordingly the smallest original dimension was performed followed by central crop. The aspect ratio of the original images was thus preserved. ResNet50 has an input size of (224×224), InceptionV3 has (299×299) and Covid-Net has (480×480). Furthermore, to mitigate the imbalance in the dataset, standard augmentations have been applied that would not change the nature of the image and would be the most natural for CXR images like - translation ($\pm$ 10%), rotation ($\pm$ 10%) and zoom ($\pm$ 15%).

4.2.2 Network Training details

For the ResNet50 [10] and InceptionV3 [27] pre-trained on ImageNet the original classification layers were replaced by an output layer of size 2 to build the binary classifiers. The training was done with Stochastic Gradient Descent with momentum (0.9 momentum coefficient) and after a grid search on {1e-6, 1e-5, 1e-4, 1e-3} learning rate of 1e-4 was chosen. The batch size of 16 was found to be the most ap-



(a) Full COVIDx dataset distribution



(b) Official COVIDx dataset distribution

Figure 7. Class distribution of COVIDx dataset.

	Normal	Covid-19
Train	6383	1697
Val	1583	437
Test	885	288

Table 1. Dataset details for Normal/Covid-19 classification. Val - Validation set

propriate with the grid search on {4, 8, 16, 32}. The loss was chosen as binary cross-entropy. All the layers in the networks were retrained on our dataset, here we used previous finding from [4] on better performances when not freezing layers, and fine-tuned for the ResNet50 for 50 epochs and InceptionV3 for 80 epochs.

Also, recently a binary version of Covid-Net called the COVID-Net-CXR-2 (Covid-Net2)[1] was made available as an open-source release that performed the Covid-19 positive/negative (normal + pneumonia) classification. This network has been fine-tuned for the task of Normal/Covid-19 classification to compare its performance to ResNet50 and InceptionV3 networks. Covid-Net2 has been fine-tuned for 10 epochs on our dataset.

4.3. Performance Evaluation

To set up a baseline, all the networks were initially trained for the binary classification of Normal vs Covid-19 classes. The best performing networks have been trained to create the multiple binary classifiers. Finally, the hierarchical classifier was created and compared to baseline Covid-Net [29]. The following sections present the results in detail.

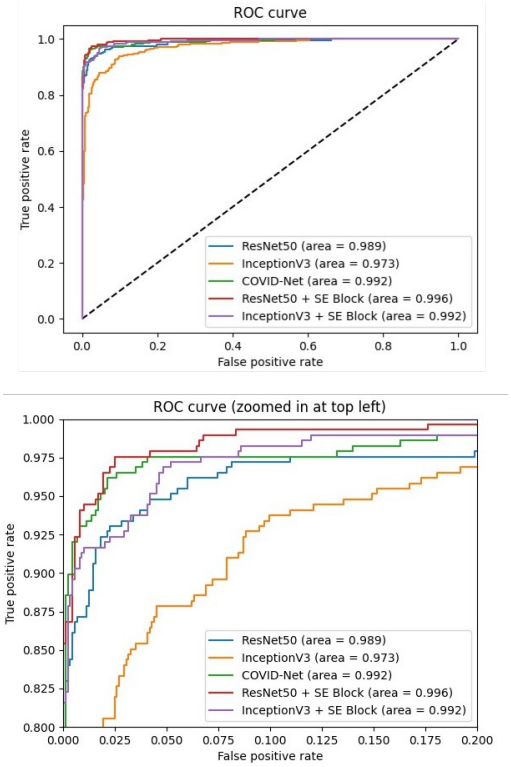


Figure 8. ROC curves and AUC of networks for Normal/Covid-19 classification. ROC: Receiver Operating Characteristics, AUC: Area Under the Curve

4.3.1 Covid-19 vs Normal Lung Classification

We have used the full COVIDx dataset to train the networks as shown in Table 1. A 90-10% train-validation split was used to tune training parameters. The results for the different networks trained for the binary classification of Covid-19 vs Normal are given in Table 2. Covid-Net performs better than the baseline networks of ResNet50 and InceptionV3 pre-trained on ImageNet achieving a better balanced accuracy (BAcc). However, InceptionV3 has a better sensitivity score of 88.5% when compared to the 84.4% of the COVID-Net for the Covid-19 class. This is essential as misclassification of the Covid-19 would have more impact than a misclassification of the normal class.

The performance of the pre-trained networks improves by adding the global attention block. InceptionV3 + SE achieves the best Balanced Accuracy (BAcc) of 93.4% and also has the highest sensitivity score of 89.6% for the Covid-19 class. Though ResNet50 + SE has the highest accuracy and sensitivity for the Normal class, the sensitivity for the Covid-19 class is lower and hence the BAcc score is low.

Architecture	Class	PPV	Sensitivity	Accuracy	BAcc
ResNet50	Normal	0.931	0.945	0.905	0.865
	Covid-19	0.822	0.785		
InceptionV3	Normal	0.962	0.932	0.921	0.909
	Covid-19	0.810	0.885		
Covid-Net	Normal	0.952	**0.999**	0.961	0.921
	Covid-19	0.996	0.844		
Resnet50 + SE Block	Normal	0.955	**0.999**	**0.963**	0.926
	Covid-19	**0.996**	0.854		
InceptionV3 + SE Block	Normal	**0.963**	0.989	0.962	**0.934**
	Covid-19	0.962	**0.896**		

Table 2. Results for per class and dataset metrics for different networks for the binary classification of Normal and COVID-19 classes. PPV - Positive Predicted Value, BAcc - Balanced Accuracy, SE - Squeeze-Excitation

The Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) values for each network is shown in Fig 8. All the networks show a similar performance. The ResNet50 + SE which achieves the highest AUC score of 0.996. Both COVID-Net and InceptionV3 + SE have a similar but slightly lower AUC of 0.992. The simple InceptionV3 network has the lowest AUC and has a lower discriminative ability when compared to the other networks.

Comparing the models it is seen that the baseline InceptionV3 has better Covid-19 sensitivity. By adding a simple attention block the performance of both the pre-trained networks improves and is better than the tailor-made COVID-Net.

4.3.2 Hierarchical Classification

Both the pre-trained networks with attention blocks show a better performance on the binary classification of Covid-19/ Normal. They have been trained to create the different classifiers proposed in Section 3. The results for the three classification tasks of 1) Normal vs Positive (Pneumonia + Covid-19), 2) Normal vs Pneumonia and 3) Covid-19 vs Pneumonia are presented in Table 3 for the two networks on the full COVIDx dataset (Fig 7(a)). InceptionV3 + SE has better BAcc of 94.3% and 91.9% for classifiers 1) and 3) when compared to ResNet50+SE for the class with the presence of disease. ResNet50 + SE has a better BAcc score for the classification of Normal/Pneumonia of 93% and a higher sensitivity of 95% for the normal class. However, the aim is to have a better prediction for the Pneumonia class (disease), for which the InceptionV3 + SE classifier has a higher sensitivity. Considering these factors, the InceptionV3 + SE network was finally chosen to build the Hierarchical classification.

The results for COVID-Net are referenced from [29] and were reported on the official COVIDx dataset (Fig. 7(b)). The results for the Inception + SE networks in Table 4 are also reported on the official COVIDx dataset. The confusion matrix for each network is shown in Fig 9. Multi-class Inception + SE has a performance lower than the baseline COVID-Net that has been designed specifically for this dataset. The network only has 84% sensitivity for Covid-19. In comparison, the hierarchical classifier comprising four Inception + SE binary classifiers shows an overall improvement over COVID-Net. Even with a lower amount of training data, we obtain better results on the same test set. From the matrix in Fig 9, the highest error Inception + SE makes is the misclassification of Covid-19 as Pneumonia. This is due to the similarity in the features of the two classes and the presence of a large number of pneumonia samples in the dataset.

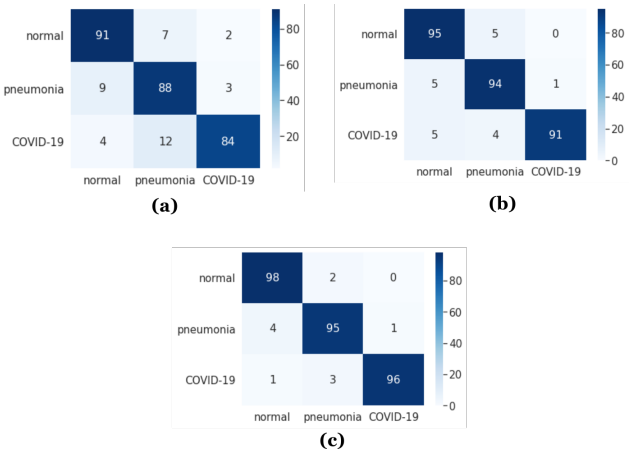


Figure 9. Confusion matrix for (a) Multi-class Inception + SE, (b) COVID-Net, (c) Multi-level binary classifiers Inception + SE, on the smaller COVIDx test set of 100 samples

5. Conclusion

In this work, we presented a hierarchical classification pipeline for the detection of Covid-19 from Chest X-Ray (CXR) images. CXR images are a low cost and easily accessible modality when compared to CT imaging. The pipeline consists of a combination of binary classifiers of

Architecture	Class	PPV	Sensitivity	Accuracy	BAcc
ResNet50 + SE Block	Normal	0.889	**0.958**	0.919	0.909
	Positive	0.954	0.880		
	Normal	0.942	**0.950**	0.935	**0.930**
	Pneumonia	0.925	0.912		
	Covid-19	0.959	0.892	0.949	0.937
	Pneumonia	0.950	0.981		
InceptionV3 + SE Block	Normal	0.935	0.933	0.923	**0.919**
	Positive	0.901	**0.904**		
	Normal	0.915	0.901	0.926	0.9215
	Pneumonia	0.934	**0.942**		
	Covid-19	0.963	**0.903**	0.956	**0.943**
	Pneumonia	0.954	**0.983**		

Table 3. Results of the metrics for the training of the ResNet50 and InceptionV3 networks with the added SE Block for the different binary classifiers of the Multi-level classification pipeline. Positive - Pneumonia + COVID-19 classes, PPV - Positive Predictive Value, BAcc - Balanced Accuracy

Architecture	Sensitivity			PPV		
	Norm	Pneu	COVID-19	Norm	Pneu	COVID-19
InceptionV3 + SE (Multi-class)	0.910	0.880	0.840	0.875	0.822	0.944
COVID-Net	0.950	0.940	0.910	0.905	0.913	0.989
Multi-Level Classifier InceptionV3 + SE (Binary Classifiers)	**0.980**	**0.950**	**0.960**	**0.951**	**0.950**	**0.990**

Table 4. Classification metrics on the official COVIDx test set of 100 samples for the networks. Norm - Normal, Pneu - Pneumonia

popular CNN networks that have been trained using cross-domain transfer learning. The networks have been modified with a low computation global attention block and the final results show an improvement from 91% to 96% for the sensitivity of the Covid-19 class when compared to the tailor-made baseline multi-class classifier network COVID-Net. In the perspective of our work, we will focus on explanations of the network decisions with methods we are developing now to make the networks trustworthy for medical practitioners.

Acknowledgment

We thank radiologist I. Saurin from Bergonié Institute Aquitania for fruitful discussions about CXR images in Covid-19 study.

References

[1] Karim Aderghal, Karim Afdel, Jenny Benois-Pineau, Gwénaëlle Catheline, and Alzheimer's Disease Neuroimaging Initiative. Improving alzheimer's stage categorization with convolutional neural network using transfer learning and different magnetic resonance imaging modalities. *Heliyon*, 6(12):e05652, 2020.

[2] Karim Aderghal, Jenny Benois-Pineau, Karim Afdel, and Gwénaëlle Catheline. Fuseme: Classification of smri images by fusion of deep cnns in 2d+ projections. In *Content-based Multimedia Indexing*, 2017.

[3] Ioannis D Apostolopoulos, Sokratis I Aznaouridis, and Mpesiana A Tzani. Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases. *Journal of Medical and Biological Engineering*, 40:462–469, 2020.

[4] Souad Chaabouni, Jenny Benois-Pineau, Francois Tison, Chokri Ben Amar, and Akka Zemmari. Prediction of visual attention with deep CNN on artificially degraded videos for studies of attention of patients with dementia. *Multim. Tools Appl.*, 76(21):22527–22546, 2017.

[5] Audrey Chung. Actualmed covid-19 chest x-ray dataset initiative, 2020.

[6] Audrey Chung. Figure 1 covid-19 chest x-ray dataset initiative, 2020.

[7] Joseph Paul Cohen, Paul Morrison, Lan Dao, Karsten Roth, Tim Q Duong, and Marzyeh Ghassemi. Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:2006.11988*, 2020.

[8] Diletta Cozzi, Marco Albanesi, Edoardo Cavigli, Chiara Moroni, Alessandra Bindi, Silvia Luvarà, Silvia Lucarini, Simone Busoni, Lorenzo Nicola Mazzoni, and Vittorio Miele. Chest x-ray in new coronavirus disease 2019 (covid-19) infection: findings and correlation with clinical outcome. *La radiologia medica*, 125:730–737, 2020.

[9] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image

database. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. Ieee, 2009.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[11] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[12] Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The lancet*, 395(10223):497–506, 2020.

[13] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 590–597, 2019.

[14] Shervin Minaee, Rahele Kafieh, Milan Sonka, Shakib Yazdani, and Ghazaleh Jamalipour Soufi. Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. *Medical image analysis*, 65:101794, 2020.

[15] Chiara Moroni, Diletta Cozzi, Marco Albanesi, Edoardo Cavigli, Alessandra Bindi, Silvia Luvarà, Simone Busoni, Lorenzo Nicola Mazzoni, Stefano Grifoni, Peiman Nazerian, et al. Chest x-ray in the emergency department during covid-19 pandemic descending phase in italy: Correlation with patients' outcome. *La radiologia medica*, 126(5):661–668, 2021.

[16] Zahra Nabizadeh-Shahre-Babak, Nader Karimi, Pejman Khadivi, Roshanak Roshandel, Ali Emami, and Shadrokh Samavi. Detection of covid-19 in x-ray images by classification of bag of visual words using neural networks. *Biomedical Signal Processing and Control*, page 102750, 2021.

[17] Radiological Society of North America. Covid-19 radiography database., 2019.

[18] Radiological Society of North America. Rsna pneumonia detection challenge, 2019.

[19] Ahmed Hamza Osman, Hani Moetque Aljahdali, Sultan Menwer Altarrazi, and Ali Ahmed. Som-lwl method for identification of covid-19 on chest x-rays. *PloS one*, 16(2):e0247176, 2021.

[20] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in biology and medicine*, 121:103792, 2020.

[21] Tuan D Pham. Classification of covid-19 chest x-rays with deep learning: new models or fine tuning? *Health Information Science and Systems*, 9(1):1–11, 2021.

[22] Md Mamunur Rahaman, Chen Li, Yudong Yao, Frank Kulwa, Mohammad Asadur Rahman, Qian Wang, Shouliang Qi, Fanjie Kong, Xuemin Zhu, and Xin Zhao. Identification of covid-19 samples from chest x-ray images using deep learning: A comparison of transfer learning approaches. *Journal of X-ray Science and Technology*, pages 1–19, 2020.

[23] Simone Schiaffino, Stefania Tritella, Andrea Cozzi, Serena Carriero, Lorenzo Blandi, Laurenzia Ferraris, and Francesco Sardanelli. Diagnostic performance of chest x-ray for covid-19 pneumonia during the sars-cov-2 pandemic in lombardy, italy. *Journal of thoracic imaging*, 35(4):W105–W106, 2020.

[24] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015.

[25] Nicola Sverzellati, Christopher J Ryerson, Gianluca Milanese, Elisabetta A Renzoni, Annalisa Volpi, Paolo Spagnolo, Francesco Bonella, Ivan Comelli, Paola Affanni, Licia Veronesi, et al. Chest x-ray or ct for covid-19 pneumonia? comparative study in a simulated triage setting. *European Respiratory Journal*, 2021.

[26] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[27] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian J. Goodfellow, and Rob Fergus. Intriguing properties of neural networks. In *ICLR (Poster)*, pages 1–10, 2014.

[28] Yu-Xing Tang, You-Bao Tang, Yifan Peng, Ke Yan, Mohammadhadi Bagheri, Bernadette A Redd, Catherine J Brandon, Zhiyong Lu, Mei Han, Jing Xiao, et al. Automated abnormality classification of chest radiographs using deep convolutional neural networks. *NPJ digital medicine*, 3(1):1–8, 2020.

[29] Linda Wang, Zhong Qiu Lin, and Alexander Wong. Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Nature Scientific Reports*, 10(1):19549, Nov 2020.

[30] Wenling Wang, Yanli Xu, Ruqin Gao, Roujian Lu, Kai Han, Guizhen Wu, and Wenjie Tan. Detection of sars-cov-2 in different types of clinical specimens. *Jama*, 323(18):1843–1844, 2020.

[31] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017.

[32] Jiong Wu, Xiaojia Wu, Wenbing Zeng, Dajing Guo, Zheng Fang, Linli Chen, Huizhe Huang, and Chuanming Li. Chest ct findings in patients with coronavirus disease 2019 and its relationship with clinical features. *Investigative radiology*, 55(5):257, 2020.

[33] Chaochao Yan, Jiawen Yao, Ruoyu Li, Zheng Xu, and Junzhou Huang. Weakly supervised deep learning for thoracic disease classification and localization on chest x-rays. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 103–110, 2018.

[34] Wenjing Yang, Arlene Sirajuddin, Xiaochun Zhang, Guanshu Liu, Zhongzhao Teng, Shihua Zhao, and Minjie Lu. The role of imaging in 2019 novel coronavirus pneumonia (covid-19). *European radiology*, pages 1–9, 2020.

[35] Dalia Yousri, Mohamed Abd Elaziz, Laith Abualigah, Diego Oliva, Mohammed AA Al-Qaness, and Ahmed A Ewees. Covid-19 x-ray images classification based on enhanced fractional-order cuckoo search optimizer using heavy-tailed distributions. *Applied Soft Computing*, 101:107052, 2021.

[36] Na Zhu, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang Zhao, Baoying Huang, Weifeng Shi, Roujian Lu, et al. A novel coronavirus from patients with pneumonia in china, 2019. *New England journal of medicine*, 2020.