CAM of RESNET18 trained on Imagenet for Covid-19 pneumonia detection

Carlo Amodeo, Andrea Benedetti, Silvia Leccabue, Isabella Poles
University of Illinois at Chicago,
{camode2, abened6, slecca2, ipoles2}@uic.edu

***Abstract*— COVID-19 disease is one of the biggest challenges of the 21st century. At the time of this work, about 150 million people have been tested positive, and more than 3 million of people have died as a result. The fighting of this virus has required, and still requires, heroism of the healthcare workers, efficiency in the research for an effective vaccination, social organization and technological solutions. As of now, the main clinical tool currently in use for the diagnosis of COVID-19 is the Reverse Transcription Polymerase Chain Reaction (RT-PCR), which is expensive, less-sensitive and requires specialized medical personnel. On the other hand, X-ray imaging is an easily accessible tool currently worldwide performed, and that could be exploited as an alternative in the COVID-19 diagnosis. This work proposes a robust technique for the automatic detection of COVID-19 and viral pneumonia from digital chest X-ray images applying the ResNet18 pre-trained deep-learning algorithm while maximizing the detection accuracy. Moreover, the Class Activation Mapping has been implemented to highlight the most relevant pixels which led to the classification of the image. The effectiveness of the proposed method was evaluated on a public chest X-ray images database, containing images labelled as COVID-19, viral pneumonia and normal images. The networks were trained to classify two different schemes: i) normal and COVID-19 pneumonia; ii) normal, viral and COVID-19 pneumonia, while 3-folds cross validation was performed to inspect the presence of sub-dataset able to better or equally generalize the problem. The classification accuracy reached the 95% level. Hence, this work shows how computer aided tools can be used to significantly improve the speed and the accuracy of the COVID-19 diagnosis, as well as reduce the work load of the healthcare workers.**

I. Introduction

Coronavirus Disease 2019 (COVID-19) is an exponential growth rate disease that has been declared a global pandemic by the World Health Organization (WHO) on 11th March 2020 and that has overloaded worldwide the healthcare systems [1]. While most of the people infected with the COVID-19 experienced mild to moderate respiratory illness, some developed a deadly pneumonia. The exact mortality, prevalence, and transmission dynamics remain somewhat ill-defined in part due to the unique challenges presented by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2 infection), such as peak infectiousness, preceding symptoms onset and a poorly understood multi-organ pathophysiology. Early diagnosis of this disease increases the chances for successful treatment of infected patients and also reduces the chances of spreading in the community for a contagious disease like COVID-19 [2]. In this direction, real-time reverse Transcriptase–Polymerase Chain Reaction (rRT-PCR) of nasopharyngeal swabs typically have been used to confirm the clinical diagnosis. However, this technique is unfortunately time-consuming, manual, laborious and with a positively rate of only 63%. Moreover, the significant shortage of its supply leads to delay in the disease prevention efforts [3]. Another technique for diagnosing lung related diseases is the chest X-ray (CXR), that is commonly available and is faster and cheaper, but signals associated with the presence of COVID-19 in the lungs can be hard to detect if investigated manually by the radiologists. Based on these considerations, fast, accurate, and rapid Artificial Intelligence (AI) models could therefore help in overcoming the lack of specialist doctors, and providing the patients with timely support. AI approaches can also help in handling some problems, such as inadequate RT-PCR test kits, and the long waiting time for test results [4]. For these reasons, the literature shows several deep machine learning techniques using chest X-ray images for classifying COVID-19, viral pneumonia and normal images [5]. However, most of these groups used rather a small dataset containing only a few COVID-19 samples. Within this context, we propose a CNN based transfer learning approach for automatic detection of COVID-19 pneumonia. An efficient pretrained Residual Neural Network, ResNet18, was used for classifying normal, viral pneumonia and COVID-19 patients chest x-rays images taken from a large publicly available database. The main contribution of this work are:

- design and implementation of an automatic pipeline for normal, viral pneumonia and COVID-19 chest x-rays image classification task using the ResNet18 deep learning strategy (Section III-B) ;
- experiments evaluation using a stratified 3-fold cross-validation (CV) scheme (Section III-E);
- the feature maps of the deeper layer of the transfer learning model are used to discriminate image regions used by the network to identify the classified category (Section III-D);.

The rest of the report is organized as follows. Section II reviews the previous studies on CNN based transfer learning approach for automatic detection of COVID-19 pneumonia. Section III provides the detailed description of the implemented COVID-19 pneumonia detection workflow, giving details about the transfer learning-based pre-trained CNN model. Section IV presents the experimental results, while conclusions are drawn in Section V.

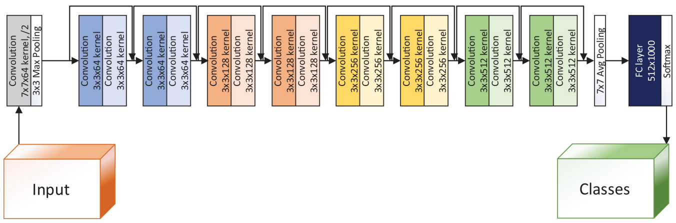


Fig. 1. The architecture of ResNet18 model.

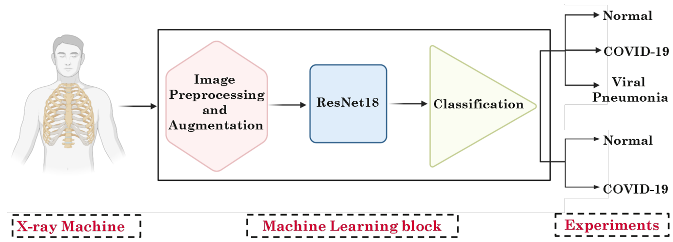


Fig. 2. Block design of the three stage pipeline implemented for the COVID-19 pneumonia detection.

II. Related Work

Artificial Intelligence-based diagnosis systems can help ease the burden on health professionals, while increasing the detection rate of COVID-19. In particular, deep learning techniques on chest X-Rays are getting popularity with the availability of the deep CNNs and the promising results it has shown in different applications. Recently, a large number of works have been carried out to detect COVID-19 using X-ray images with the help of AI models. Ioannis et al. [6] reported 96.78 % accuracy for COVID-19 from bacterial pneumonia and normal X-rays in a dataset of 1427 X-rays. Similarly, Abbas et al. [7] reported an accuracy of 95.12 % for COVID-19 classification from COVID-19, normal and SARS CXR's using their pre-trained CNN model (DeTraC Decompose, Transfer and Compose) with a small database of 196 X-ray images. Minaee et al. [8] reported a specificity and sensitivity of 90% and 97 % respectively using ChexPert dataset [9]. Ashfar et al. [10] reported an accuracy of 95.7 % using a Capsule Networks, called COVID-CAPS rather than a conventional CNN to deal with a smaller dataset. These results showed the potential of using CNN to distinguish COVID- 19 from other lung diseases using CXR images. However, most of these groups used rather a small dataset containing only a few COVID-19 samples. This makes it difficult to generalize their results reported in these articles and cannot guarantee that the reported performance will retain when these models will be tested on a larger dataset. A rigorous experiment on a large database of COVID-19 and non-COVID-19 classes are very few and missing in case of transfer learning approach. However Chowdhury et al. [11] achieved good results training, validating and testing three shallow networks (MobileNetv2, SqueezeNet and ResNet18) with other five deep networks (Inceptionv3, ResNet101, CheXNet, VGG19 and DenseNet201) to automatically detect COVID-19 pneumonia from chest x-rays images taken from a large publicly available database. Following the line of this last study and exploiting the same dataset, we propose an automatic pipeline to perform two classification tasks: one classification model was trained to classify COVID-19 and normal X-ray images, while the other was trained to classify normal, viral pneumonia and COVID-19 pneumonia images using the ResNet18 deep learning strategy. In addition, in order to understand for a particular category the discriminative image regions used by the CNN to identify that category, class activation maps were extracted.

III. Proposed framework

This Section presents an in-depth description of the proposed workflow, focusing on the image processing and data augmentation, devised to deal with the CNN input requirements, on the CNN model selection to perform the classification task and finally on the class activation maps' extraction in order to have a deeper understanding about the CNN feature extraction strategy. As shown in Figure 2, the images are given in input to the machine learning block, which consists on the data processing and on the CNN training, while the two classification tasks conclude the framework flow.

A. Dataset description

In this study, Posterior-to-Anterior/Anterior-to-Posterior (AP/PA) image of chest X-ray was used, as this view of radiography is widely used by radiologist in clinical diagnosis. This study has been made possible thanks to a team of researchers from Qatar University (Doha, Qatar) and the University of Dhaka, along with their co-workers from Pakistan and Malaysia who, in collaboration with medical doctors, have created a database of chest X-ray images for COVID-19 positive cases along with normal and viral pneumonia images [11].

This COVID-19, normal, and other lung infection dataset has been released in stages, is continuously being updated as soon as new AP/PA images of chest X-ray are available to its authors and is provided to the global research community to apply recent advances in deep learning, image processing and other AI techniques to generate new insights in support of the ongoing fight against this infectious disease.

It has been developed starting from the database of COVID-19 X-ray images from the Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 DATABASE, Novel Corona Virus 2019 Dataset developed by Joseph Paul Cohen and Paul Morrison, and Lan Dao in GitHub and images extracted from other 43 different publications. Regarding normal and viral pneumonia images, they have been adopted from the chest X-Ray images (pneumonia) database.

In summary, the used dataset is composed by 15153 portable network graphics (PNG) images whose size is 299x299 pixels and which are labelled as "COVID-19", "NORMAL" or "VIRAL" (Figure 3). There is not a great unbalancing between the different classes, as the number of images per labels are

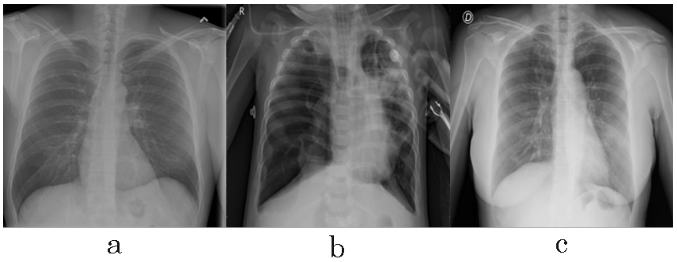


Fig. 3. Sample X-ray image from the dataset: normal X-ray image (a), viral X-ray image (b), and COVID-19 X-ray image (c).

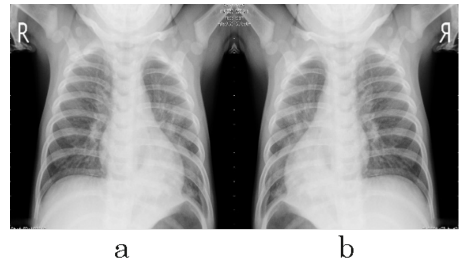


Fig. 4. Original image (a), result of the image augmentation (b).

respectively 3616 COVID-19, 10192 normal and 1345 viral pneumonia.

B. CNN Model Selection

When deeper networks are able to start converging, a degradation problem has been exposed: with the network depth increasing, accuracy gets saturated and then degrades rapidly. Unexpectedly, such degradation is not caused by overfitting, and adding more layers to a suitably deep model leads to higher training error.

In general, ResNet18 is one of the variants of the Residual Network (in short ResNet) in which vanishing gradient and degradation problem are solved. In addition, this class of algorithms learns from residuals instead of features.

Concerning the architecture, ResNet18 is characterized by 18 layers and is mostly composed by 3x3 filters, as Figure 1 shows. A down sampling has been directly performed by convolutional layers that have a stride of 2.

The network ends with a global average pooling layer and a 1000-way fully-connected layer with Softmax.

Furthermore, shortcut connections have been inserted that turn the network into its counterpart residual version.

This pre-trained CNN model was trained using Stochastic Gradient Descent (SGD) with momentum optimizer with learning rate $\alpha = 10^{-5}$, momentum update $\beta = 0.1$ and mini-batch size of 6 images with 20 back propagation epochs, using Adam optimizer.

In addition, given that the dataset is composed by 3 classes as mentioned in III-A and ResNet18 is a pre-trained deep CNN with 1000 classes as output features, changing the last fully connected layer has been an essential step in order to achieve the final objective of the project.

Two different experiments were carried out in this study, that are a two-class image classification using model trained with images augmentation, and then, a three-class image classification using models trained with image augmentation, as well.

In both cases, cross entropy has been applied as loss function and as we have mentioned before, the chosen optimizer is Adam optimizer.

The only two classified classes in the first mentioned experiment are just COVID-19 and normal. This because it could be plausible that in a real life-clinical application, the worst scenario would be the one in which we miss-classify a normal case with a COVID-19 one, and viceversa; hence, we wanted to be sure that the classification would be performed properly in that scenario.

C. Preprocessing and Image Augmentation

Chest X-ray images were only resized before applying them as input to the network: input requirements for different CNNs are different and, in particular, ResNet18 requires input data whose size is 224x224. Every single image has been normalized, according to the pre-trained model standards.

In order to increase the robustness of the model, data augmentation techniques have been used. Since the dataset is not an unbalanced one, it has been decided to augment all the classes in the same way. Using a simple horizontal flipping function, and hence increasing the quantity of available images of a factor whose value is less than two, has been enough in order to obtain satisfactory results in terms of accuracy of the model.

D. Class Activation Map

Zhou et al [12] reported that the convolutional units of various layers of convolutional neural networks actually behave as object detectors even if no supervision on the location of the object was provided.

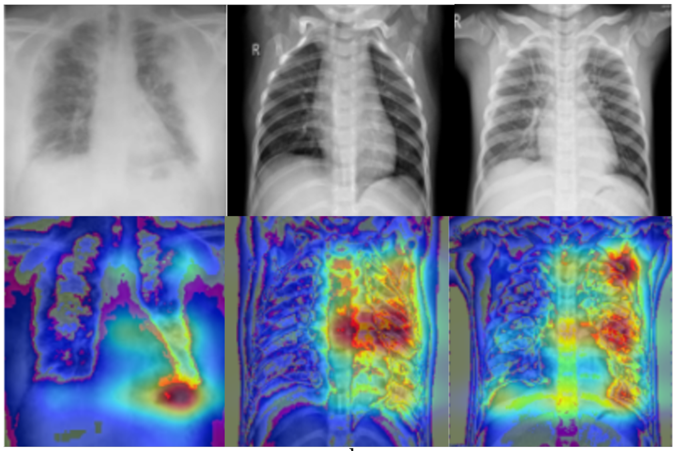
Firstly, in order to achieve the capability to localize object in the convolutional layers, the use of a global average pooling becomes fundamental, because a fully-connected layers used for classification would lead to lose this ability.

Global average pooling acts as a structural regularizer that prevents overfitting during training. Furthermore, it is performed before the final output layer on the convolutional feature maps and use those as features for a fully-connected layer that produces the desired output.

This simple connectivity structure belongs to a technique called Class Activation Mapping (CAM) and for a particular category indicates the discriminative image regions used by the CNN to identify that specific category.

It is important to state that we can identify the importance of the image regions by projecting back the weights of the output layer on to the convolutional feature maps. [13]

Given that the activation map can take different range of values, in this specific project a normalization between 0 and 1 has been applied. Finally, the strongest activation channels from the COVID-19, normal and viral pneumonia X-ray images were identified and compared with the original images (Figure 5).



a b c

Fig. 5. On the top row original images, on the bottom row the Class Activation Maps, respectively for a COVID-19 image (a), normal image (b), viral pneumonia image (c).

E. Random Sub-sample Cross Validation

A random sub-sample cross-validation has been used with the aim of inspecting if there is any sub-dataset able to better or equally generalize the problem. Moreover, this type of algorithm may be useful to speed up the process of training and try to minimize the reward over time by the network. The iteration of a random sub-sampling of 900 images equally divided per labels has been performed 3 times. Furthermore, both the experiments were evaluated using the same cross-validation scheme with a ratio of 90% for training and 10% for the test splits.

F. Performance evaluation metric

The performance of the networks was evaluated using accuracy (*Acc*) as performance metric in both classification schemes. Accuracy can be a reliable and significant performance metric in this specific project, because the distribution of examples in the testing dataset across the classes is balanced. Formally, this parameter is computed as:

$$Accuracy_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{1}$$

where i corresponds to COVID-19 and normal for the two class problem; COVID-19, normal and viral pneumonia for the three class problem.

IV. RESULTS

This Section presents an evaluation of the reached accuracy performance respectively for the two classification tasks, the cross validation experiment and the class activation map interpretation. The Accuracy rate (*Acc*) is analyzed to provide a quantitative assessment of the performance of the developed method.

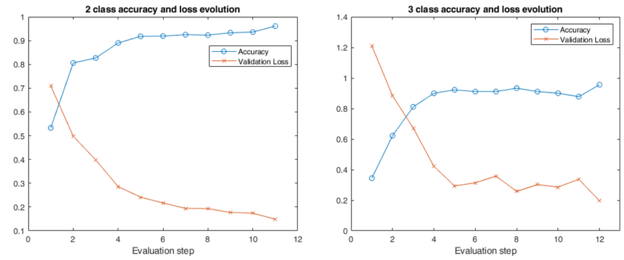


Fig. 6. On the left accuracy and validation loss plot over iterations for the 2 class problem; on the right accuracy and validation loss plot over iterations for the 3 class problem.

TABLE I
INFERENCE RESULTS ON A SINGLE IMAGE FOR THE 2 CLASS PROBLEM
N: NORMAL, C: COVID-19

Input label	**Predicted label**	**Probabilities [N, C]**
N	N	0.98, 0.02

A. Experimental results - Comparison 2 class problem and 3 class problem

The 95% accuracy was chosen as an evaluation metric for mainly two reasons: the testing subsets were balanced in 30 samples per labels and this proves the suitability of this metrics, and the 95% value has been chosen from the literature as a saturation threshold, since it is the used value for most state of the art approach in time-iteration performance evaluation. Figure 6 shows that the efforts taken by the model to reach the threshold of accuracy was remarkably less in the two labels classification, justified by less information needed to be learned by the network at each iteration, even though the 3 class problem performance are acceptable as well. The 3 classes problem solution clearly leads to a more versatile and robust model, but it has a heavier computational cost. Focusing on the Table I and II we can see the output obtained when dealing with a single image used as an input for both classification tasks: the output is an array of probabilities, telling us the most probable class to which that image belongs. It is possible to observe how the algorithm succeeded in inferring the right label for the input image.

B. Experimental results - Cross validation 2 class problem and 3 class problem

With the objective previously mentioned in Section III-E a random sub-sample cross-validation is used, iterating 3 times a random sub-sampling of 900 images equally divided per labels. The results in Figure 7 show that both in the 2 and the 3 class problem statement, the time consumed by the model to reach a 95% accuracy was significantly reduced in the sub-dataset of the cross-validation. Even though the number of iterations usually increase the cumulative time to reach our target performance, the results state that the overall timing performance improves.

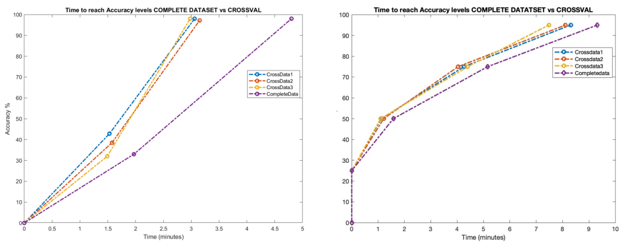


Fig. 7. On the left accuracy plot over time for the 2 class problem; on the right accuracy plot over time for the 3 class problem. Both plots show the comparison between the performance on the complete dataset and during the cross validation task.

TABLE II
INFERENCE RESULTS ON A SINGLE IMAGE FOR THE 3 CLASS PROBLEM
N: NORMAL, C: COVID-19, VP: VIRAL PNEUMONIA

Input label	Predicted label	Probabilities [N, C, VP]
N	N	0.78, 0.11, 0.11

C. Class Activation Maps

The CAM map obtained from the second-last layer was extracted in order to visually perceive by intuition which are the "key part" of the image used to classify the image itself by the deep network. From 5 is possible to highlight how the more weighted areas taken in consideration for classification are reasonably the one associated with the area of lungs' lesions.

V. CONCLUSIONS AND DISCUSSIONS

This study has introduced an high performance automated and low-cost framework for evaluating different type of pneumonia, but moreover to provide an accurate diagnosis of COVID-19 based on chest x-ray images. The pre-processing needed to feed the model is nearly in-existent and this allows an easy-to use pipeline for health-care-provider. Experimental results, focused on the *Acc* values, revealed that the proposed method is accurate in distinguish COVID-19 chest x-rays from normal and viral ones. Any future development would be to try to understand how deep-learning machines can infer so well in fields that are difficult to understand even for human experts, gaining interpretability of results that would lead to remarkable insights in such diagnostic science as essential as the the correct usage of these technologies.

REFERENCES

[1] S. A. Harmon, T. H. Sanford, S. Xu, E. B. Turkbey, H. Roth, Z. Xu, D. Yang, A. Myronenko, V. Anderson, A. Amalou *et al.*, "Artificial intelligence for the detection of covid-19 pneumonia on chest ct using multinational datasets," *Nature communications*, vol. 11, no. 1, pp. 1–7, 2020.

[2] X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan *et al.*, "Temporal dynamics in viral shedding and transmissibility of covid-19," *Nature medicine*, vol. 26, no. 5, pp. 672–675, 2020.

[3] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan, "Detection of sars-cov-2 in different types of clinical specimens," *Jama*, vol. 323, no. 18, pp. 1843–1844, 2020.

[4] R. A. Al-Falluji, Z. D. Katheeth, and B. Alathari, "Automatic detection of covid-19 using chest x-ray images and modified resnet18-based convolution neural networks," *Computers, Materials, & Continua*, pp. 1301–1313, 2021.

[5] T. T. Nguyen, "Artificial intelligence in the battle against coronavirus (covid-19): a survey and future research directions," *arXiv preprint arXiv:2008.07343*, 2020.

[6] 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, 2020.

[7] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, 2021.

[8] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning," *Medical image analysis*, vol. 65, p. 101794, 2020.

[9] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya *et al.*, "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 590–597.

[10] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images," *Pattern Recognition Letters*, vol. 138, pp. 638–643, 2020.

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

[12] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Object detectors emerge in deep scene cnns," *arXiv preprint arXiv:1412.6856*, 2014.

[13] ——, "Learning deep features for discriminative localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2921–2929.