

Understanding Automatic COVID-19 Classification using Chest X-ray images

Pierangela Bruno¹[0000-0002-0832-0151], Cinzia Marte¹[0000-0003-3920-8186], and
Francesco Calimeri¹[0000-0002-0866-0834]

Department of Mathematics and Computer Science,
University of Calabria, Rende, Italy
{bruno,marte,calimeri}@mat.unical.it

Abstract. The COVID-19 disease caused by the SARS-CoV-2 virus first appeared in Wuhan, China, and is considered a serious disease due to its high permeability, and contagiousness. The similarity of COVID-19 disease with other lung infections, along with its high spreading rate, makes the diagnosis difficult. Solutions based on machine learning techniques achieved relevant results in identifying the correct disease and providing early diagnosis, and can hence provide significant clinical decision support; however, such approaches suffer from the lack of proper means for interpreting the choices made by the models, especially in case of deep learning ones. With the aim to improve interpretability and explainability in the process of making qualified decisions, we designed a system that allows a partial opening of this black box by means of proper investigations on the rationale behind the decisions. We tested our approach over artificial neural networks trained for multiple classification based on Chest X-ray images; our tool analyzed the internal processes performed by the networks during the classification tasks to identify the most important elements involved in the training process that influence the network's decisions. We report the results of an experimental analysis aimed at assessing the viability of the proposed approach.

Keywords: GradCAM · Chest X-ray images · Convolutional Neural Networks

1 Introduction

The Novel Coronavirus, that reportedly started to infect human individuals at the end of 2019, rapidly caused a pandemic, as the infection can spread quickly from individual to individual in the community [16]. Signs of infection include respiratory symptoms, fever, cough and dyspnea. In more serious cases, the infection can cause Pneumonia, severe acute respiratory syndrome, septic shock, multi-organ failure, and death [15, 13].

Early and automatic diagnoses are relevant to control the epidemic, paving the way to timely referral of patients to quarantine, rapid intubation of serious

Copyright ©2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0)

cases in specialized hospitals, and monitoring of the spread of the disease. Since the disease heavily affects human lungs, analyzing Chest X-ray images of the lungs may prove to be a powerful tool for disease investigation. Several methods have been proposed in the literature in order to perform disease classification from Chest X-ray images, especially based on deep learning approaches [29, 9, 4]. Notably, in this context, solutions featuring interpretability and explainability approaches can significantly help at improving disease classification and providing context-aware assistance and understanding. Indeed, interpreting the decision-making processes of neural networks can be of great help at enhancing the diagnostic capabilities and providing direct patient- and process-specific support to diagnosis and surgical tool detection. However, interpretability and explainability represent critical points in approaches based on deep learning models, that achieved great results in disease classification.

In this work, we investigate the use of convolutional neural networks (CNNs) with the aim to perform multiple-disease classification from Chest X-ray images. Diseases that are a matter of concern for our experiments are COVID-19, Viral Pneumonia and Streptococcus Pneumonia. Notably, although these diseases are characterized by pulmonary inflammation caused by different pathogens, Streptococcus Pneumonia and Viral Pneumonia have similar clinical symptoms to COVID-19 [24, 28], such as fever, chills, cough, and dyspnea. The symptom-based similarity among diseases is a critical factor that could affect a proper diagnosis and treatment plan. Moreover, we include in our experiments Healthy patients to learn how they differ from symptomatic patients.

We analyze the CNNs-based model to identify the mechanisms and the motivations steering neural networks decisions in classification task. In particular, we use gradient visualization techniques to produce coarse localization maps highlighting the image regions most likely to be referred to by the model when the classification decision is taken. The highlighted areas are then used to discover (i) patterns in Chest X-ray images related to a specific disease, and (ii) correlation between these areas and classification accuracy, by analyzing a possible performance worsening after their removal.

The remainder of the paper is structured as follows. We first briefly report on related work in Section 2; in Section 3 we then provide a detailed description of our approach, that has been assessed via a careful experimental activity, which is discussed in Section 4; we analyze and discuss results in Section 5, eventually drawing our conclusions in Section 6.

2 Related work

In this section we present state-of-the-art methods used to (i) perform disease classification through Chest X-ray images and (ii) provide interpretability and explainability of the rationale behind the decisions performed.

Disease Classification. Deep learning-based models recently achieved promising results in image-based disease classification. These models, such as CNNs [14,

6, 20, 25], are proven to be appropriate and effective when compared to conventional methods; indeed, CNNs currently represent the most widely used method for image processing. Abbas et al. [1] proposed a deep learning approach (DeTraC) to perform disease classification using X-ray images. The approach was used to distinguish COVID-19 X-ray images from normal ones, achieving an accuracy of 95.12%. An improvement in terms of binary classification accuracy was presented by Ozturk et al. [17]. The authors proposed a deep learning model (DarkCovidNet) for automatic diagnosis of COVID-19 based on Chest X-ray images. They both performed a binary and multiclass classification, dealing with patients with COVID-19, no-findings and Pneumonia. The accuracy achieved is of 98.08% and 87.02%, respectively. Similarly, Wang et al. [22] proposed a deep learning-based approach (COVID-Net) to detect distinctive abnormalities in Chest X-ray images among patients with non-COVID-19 viral infections, bacterial infections, and healthy patients, achieving an overall accuracy of 92.6%. All the approaches showed limitations related to low number of image samples and imprecise localization on the chest region. More accurate localization of model’s prediction was proposed by Mangal et al. [11] and Haghanifar et al. [7]. The authors proposed a deep learning-based approach to classify COVID-19 patients from others/normal ones. They also generated saliency maps to show the classification score obtained during the prediction and to validate the results.

Explainability of deep learning model. In the last year, attempts at understanding neural networks decision-making have raised a lot of interest in the scientific community. Several approaches have been proposed to visualize the behavior of a CNN by sampling image patches that maximize the activation of hidden units [26], and by backpropagation to identify or generate salient image features [10]. Other researchers were trying to solve this problem by explaining neural network decisions by generating informative heatmaps such as Gradient-weighted Class Activation Mapping (GradCAM) [19, 3], or through layer-wise relevance propagation [2]. However, these methods present some limitations; indeed, the generated heatmaps were basically qualitative, and not informative enough to specify which concepts have been detected. An improvement was provided using semantically explanation from visual representation [27] to decompose the evidence for a prediction for image classification into semantically interpretable components, each with an identified purpose, a heatmap, and a ranked contribution.

In this work, we propose the use of Deep Learning approach to perform multiple disease classification using Chest X-ray. Additionally, we take advantage of a novel technique for analyzing the internal processes and the decision performed by a neural network during the training phase.

3 Proposed Approach

3.1 Classification

Considering that other diseases appear similar to COVID-19 Pneumonia, including other Coronavirus infections and community-acquired Pneumonia such

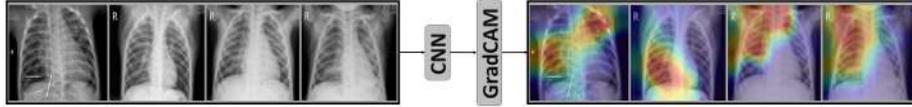


Fig. 1: Workflow of the proposed framework. Chest X-ray images are used to train the CNN. The last convolution layer of the CNN is used as input of the GradCAM approach to provide the corresponding visual explanations (i.e the regions of input that are “important” for classification capability).

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201
Conv	112×112	7×7 conv, stride 2		
Pooling	56×56	3×3 max pool, stride 2		
DB (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
TL (1)	56×56 28×28	1×1 conv		
		2×2 average pool, stride 2		
DB (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
TL (2)	28×28 14×14	1×1 conv		
		2×2 average pool, stride 2		
DB (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
TL (3)	14×14 7×7	1×1 conv		
		2×2 average pool, stride 2		
DB (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$
CL	4×1	7×7 global average pool		
		4D fully-connected, softmax		

Table 1: Architecture of the networks DenseNet-121, DenseNet-169 and DenseNet-201. More in detail, Conv stands for convolution, DB for Dense Block, TL for Transition Layer, CL for Classification Layer.

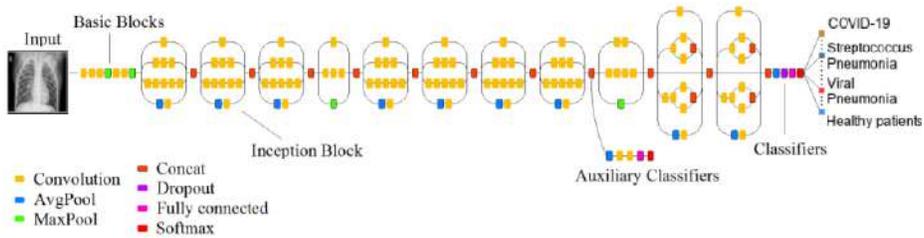


Fig. 2: Architecture of the network Inception-v3.

as Streptococcus, the distinction between these is extremely important and necessary, especially during a pandemic. Therefore, our purpose is to automatically identify the “correct” disease in Chest X-ray images.

In order to achieve this goal, we train CNN to classify patients according to 3 similar-based symptoms diseases (i.e., COVID-19, Viral Pneumonia, Streptococcus Pneumonia) and Healthy patients.

The herein proposed approach, illustrated in Fig. 1, is based on: (i) Multiple-disease classification using CNNs, and (ii) Visual Explanations using GradCAM to indicate the discriminative image regions used by the CNN.

In order to classify patients, we used and compared the results of four neural networks chosen on the basis of the good performance obtained on the *ImageNet* data set over several competitions [18]. We make use of DenseNet 121, DenseNet 169, DenseNet 201 and Inception V3 [21].

DenseNet networks [8] are made of dense blocks, as shown in Table 1, where for each layer the inputs are the feature maps of all the previous layers with the aim to improve the information flow on 224×224 input images. More in detail, for convolutional layers with kernel size 3×3 , each side of the inputs is zero-padded by one pixel to keep the feature-map size fixed. The layers between two contiguous dense blocks are referred as transition layers for convolution and pooling, which contain 1×1 convolution and 2×2 average pooling. A 1×1 convolution is introduced as a bottleneck layer before each 3×3 convolution to reduce the number of input feature-maps, and thus to improve computational efficiency. At the end of the last dense block, a global average pooling and a softmax classifier are applied.

The structure of Inception-v3 is shown in Fig. 2. The Inception modules (Inception A, Inception B and Inception C) are well-designed convolution modules that can both generate discriminatory features and reduce the number of parameters. Each Inception module is composed of several convolutional layers and pooling layers in parallel. The network is composed of 3 Inception A modules, 5 Inception B modules, and 2 Inception C modules that are stacked in series. The input size used is 224×224 and, after the Inception modules and convolutional layers, the feature map dimensions were 5×5 with 2.048 channels. Afterwards, we added 3 fully connected layers to the end of the Inception modules, and, finally, a softmax layer was added as a classifier outputting a probability for each class, and the one with the highest probability was chosen as the predicted class.

3.2 Visual Explanations

We used GradCAM to identify visual features in the input able to explain result process achieved during the multiple classification. The overall structure of GradCAM is showed in Fig. 3. In particular, it uses the gradient information flowing into the last convolutional layer of the CNN to assign importance values to each neuron. GradCAM is applied to a trained neural network with fixed weights. Given a class of interest c , let y^c the raw output of the neural network,

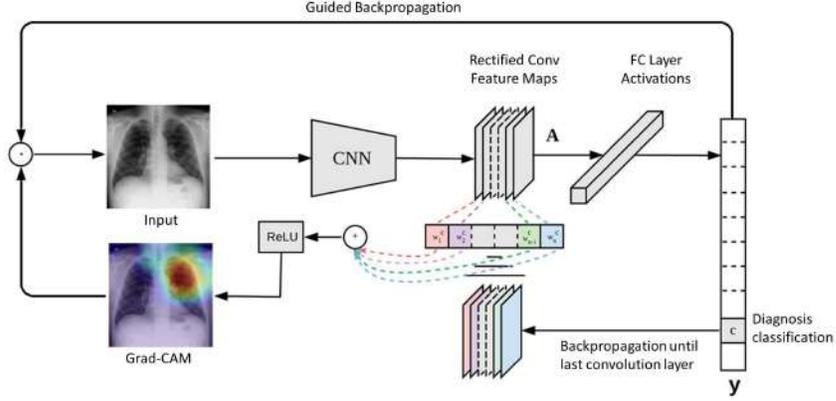


Fig. 3: An example of GradCAM structure. Given an image and a category ("Diagnosis c ") as input, we forward-propagate the image through the model to obtain the raw class scores before softmax. The gradients are set to zero for all classes, except for desired class ("Diagnosis c "), which is set to 1. This signal is then backpropagated to the rectified convolutional feature map (A) of interest, where we can compute the coarse GradCAM localization (blue heatmap).

that is, the value obtained before the application of softmax used to transform the raw score into a probability. GradCAM performs the following three steps:

1. **Compute Gradient** of y_c w.r.t. feature maps activation A^k , for any arbitrary k , of a convolutional layer (i.e., $\frac{\delta y_c}{\delta A^k}$). This gradient value depends on the input image chosen; indeed, the input image determines both the feature maps A^k and the final class score y^c that is produced.
2. **Calculate Alphas by Averaging Gradients** over the width dimension (indexed by i) and the height dimension (indexed by j) to obtain neuron importance weights α_k^c , as follows:

$$\alpha_k^c = \frac{1}{Z} \overbrace{\sum_i \sum_j}^{\text{global average pooling}} \underbrace{\frac{\delta y^c}{\delta A_{i,j}^k}}_{\text{gradients via backprop}},$$

where Z is a constant (i.e., number of pixels in the activation map).

3. **Calculate Final GradCAM Heatmap** by performing a weighted combination of the feature map activations A^k as follows:

$$L_{GradCAM}^c = ReLU \left(\underbrace{\sum_k \alpha_k^c A^k}_{\text{linear combination}} \right),$$

where α_k^c is a different weight for each k , and $ReLU$ is the Rectified Linear Unit operation used to emphasize only the positive values and to convert the negative values into 0.

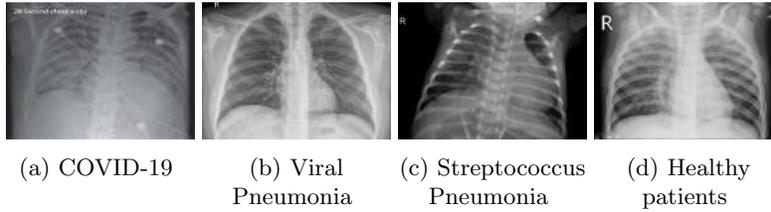


Fig. 4: Example of frontal-view chest X-ray images for the treated pathologies.

4 Experimental Protocol

We describe next the setting of the experimental analysis performed in order to assess the viability of our approach.

4.1 Dataset description

For the experimental analysis we used datasets provided by Cohen et al. [5] and Wang et al. [23]. The datasets consist of several X-ray extracted from various online publications and websites. Examples of X-ray images are shown in Fig. 4.

The datasets specifically include images of COVID-19 cases along with others, such as, for instance, Viral and Bacterial Pneumonia based images.

In particular, we considered only 4 specific categories distributed as follows:

1. COVID-19, counting 434 patients
2. Viral Pneumonia, counting 1337 patients
3. Streptococcus Pneumonia, counting 1400 patients
4. Healthy patients, counting 1341 patients

In order to obtain a valid classification and avoid majority class selection, we properly made use of data augmentation techniques to over-sample imbalanced data and obtain an equal number of samples in abundant class. More specifically, we performed:

- Translating medical images: shift the region of interest with respect to the center of the training images;
- Rotating medical images: rotate the training images by a random amount of degrees;
- Flipping medical images: use randomized flipping, through which the image information is mirrored horizontally or vertically.

4.2 Training phase

The dataset was split into training (80%) and testing (20%) sets; the 20% of the training set is used as validation set, in order to monitor the training process and prevent overfitting.

All experiments have been performed on a machine equipped with a 12 x86 64 Intel(R) Core(TM) CPUs @3.50GHz, running GNU/Linux Debian 7 and using

CUDA compilation tools, release 7.5, V 7.5.17 NVIDIA Corporation GM 204 on GeForce GTX 970.

Fine-tuning For the training phase we performed hyperparameters optimization. DenseNet 169 was trained with both optimizers Adam and SGD and for each optimizer 7 learning rate were tried. The best performance is obtained with the following configuration, trained for 300 epochs: Adam optimizer, learning rate 10^{-5} , batch size 16, and binary cross-entropy as loss function.

The configuration of networks was modified in terms of the number of nodes or levels to optimize the performance. We empirically changed the number of layers and we trimmed network size by pruning nodes to improve computational performance and identify those nodes which would not noticeably affect network performance. However, since we performed the experiments using well-know networks already optimized, we achieved the best performance using the standard configuration as originally proposed by respective authors.

We performed 10-fold cross-validation in order to choose the parameter value that gives the lowest cross-validation average error; experiments were performed on the very same machine with the same configuration of the other approaches.

4.3 Performance Metrics

We assessed the effectiveness of our approach by measuring Area Under the Curve (AUC) and Recall, especially focusing on the last one; indeed, in this context, the most important thing is to minimize False Negatives (i.e., disease is present but is not identified).

Let TP be a True Positive, TN a True Negative, FN a False Negative, and FP a False Positive, a ROC curve is a plot of true positive fractions ($S_e = \frac{TP}{TP+FN}$) versus false positive fractions ($S_p = 1 - \frac{TN}{TN+FP}$) by varying the threshold on the probability map. Closer a curve approaches the top left corner, then better is the performance of the system.

The Area Under the Curve (AUC), which is 1 for a perfect system, is a single measure to quantify this behavior [12].

Recall ($Rec = \frac{TP}{TP+FN}$) considers prediction accuracy among only actual positives and explain how correct our prediction is among all people.

5 Results and Discussion

Table 2 and Table 3 report classification results after 10-fold cross-validation for all datasets in terms of Recall and AUC, respectively. Even though promising results are achieved in all DenseNet-based experiments, DenseNet 169 shows the most efficient architecture: it reports AUC mean value of 0.95 and Recall mean value of 0.90 over all the classes: hence, it was the one selected for the study.

The herein proposed approach achieves promisingly results; in particular, DenseNet 169 achieves the best performance on COVID-19 dataset (i.e., Recall mean value: 0.99 and AUC: 0.99), instead, it decreases on Viral Pneumonia dataset (i.e., Recall mean value: 0.83 and AUC: 0.92), and in classifying Healthy patients (i.e., Recall mean value of 0.80 and AUC: 0.89).

Table 2: Validation Recall (and standard deviation) for the 4 tested neural networks after 10-fold cross-validation for each dataset. Most significant results here highlighted.

DATASET	DenseNet 121	DenseNet 169	DenseNet 201	Inception V3
COVID-19	0.98 (0.01)	0.99 (0.00)	0.96 (0.01)	0.94 (0.01)
Viral Pneumonia	0.79 (0.07)	0.83 (0.05)	0.80 (0.08)	0.77 (0.08)
Streptococcus Pneumonia	0.95 (0.01)	0.96 (0.01)	0.90 (0.01)	0.88 (0.01)
Healthy patients	0.78 (0.04)	0.80 (0.04)	0.82 (0.03)	0.84 (0.02)

Table 3: AUC values for the 4 tested neural networks after 10-fold cross-validation for each dataset. Most significant results here highlighted.

DATASET	DenseNet 121	DenseNet 169	DenseNet 201	Inception V3
COVID-19	0.99	0.99	0.99	0.97
Viral Pneumonia	0.91	0.92	0.91	0.88
Streptococcus Pneumonia	0.97	0.98	0.97	0.94
Healthy patients	0.87	0.89	0.88	0.82

Analyzing the results, we see that Viral Pneumonia is often confused with Streptococcus Pneumonia, due to overlapping imaging characteristics and thus resulting in a Recall mean value always less than 0.90 in all experiments performed. It is worth noting that, in general, the extraction of CT scan images from published articles, rather than from actual sources, might lessen image quality, thus affecting performance of the machine learning model.

A visual inspection of the GradCAM results confirms the quality of the model; indeed, our model exhibits strong classification criteria in the Chest region (see Fig. 5). In particular, red areas refer to the parts where the attention is strong, while blue areas refer to weaker attention. In general, the warmer the color, the more important are the features highlighted for the network.

Moreover, in order to confirm that the identified portions are actually significant, for each dataset we selected and removed the 40% of highlighted elements; this threshold was selected empirically after several experiments. A substantial decrease of Recall (on average around 10%) is shown using COVID-19, Viral Pneumonia and Streptococcus Pneumonia (i.e., p-value < 0.05 for paired t-test computed before and after images cutting); no statistical changes are shown using dataset of Healthy patients. This result suggests that GradCAM is able to identify the important elements involved in the training process and, consequently, responsibility for this diminishment is due to images cutting by which we removed the peculiar characteristic of the disease.

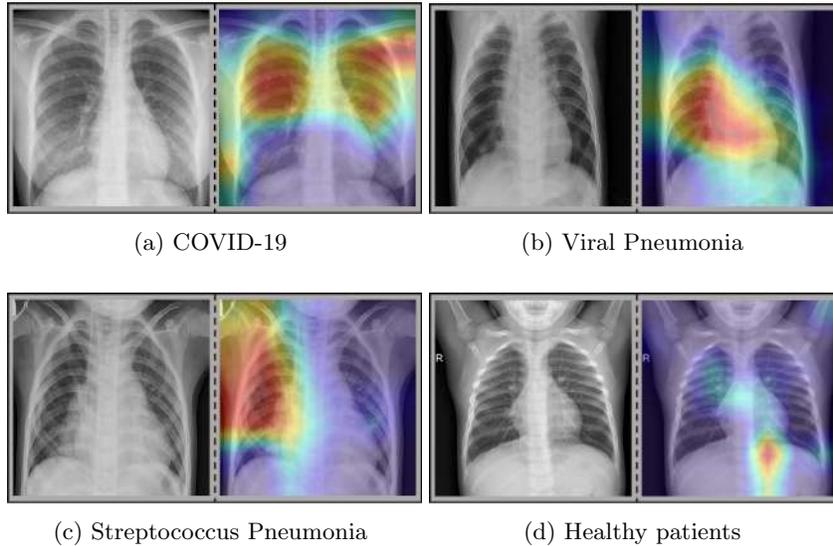


Fig. 5: Visual example of achieved results. For each diagnostic class, we show raw Chest X-ray image (left) and GradCAM result (right). Images on the right sides highlight the most important areas involved in the classification process.

6 Conclusion

In this work we exploit the use of CNNs and visual explanation techniques to estimate diagnosis using Chest X-ray and to analyze the internal processes performed by a neural network during the training phase with the aim of improving explainability in the process of making qualified decisions. Basically, we try to identify the most important regions that influence the network’s decisions.

We fine-tuned the approach by means of accurate experimental activities; in particular, we classified four different disease datasets, and four different CNNs for the classification.

Experimental results show that our proposal is robust and it is able to identify specific regions that are crucial in the neural network decision-making process, thus improving explainability. Indeed, classification accuracy is lower when highlighted regions are removed from the input images; this suggests the importance of these areas in disease classification and the possibility to consider the set of elements identified as potential disease markers.

In context where early and accurate medical diagnosis of specific pathologies are essential, our method proves that visual explanation method combined with machine learning techniques can be used to provide solid disease classifications and automatically discover new bio-markers by interpreting network decisions.

As future work is concerned, we plan to investigate misclassification errors and improve the generalization capability of the model. Our efforts will also include the interaction with physicians, so that proper medical expertise can be

used to judge and better assess the quality of the regions highlighted by the proposed approach.

References

1. Abbas, A., Abdelsamea, M.M., Gaber, M.M.: Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. arXiv preprint arXiv:2003.13815 (2020)
2. Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.R., Samek, W.: On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS one* **10**(7), e0130140 (2015)
3. Bruno, P., Calimeri, F., Kitanidis, A.S., De Momi, E.: Understanding automatic diagnosis and classification processes with data visualization. In: 2020 IEEE International Conference on Human-Machine Systems (ICHMS). pp. 1–6. IEEE (2020)
4. Bullock, J., Cuesta-Lázaro, C., Quera-Bofarull, A.: Xnet: A convolutional neural network (cnn) implementation for medical x-ray image segmentation suitable for small datasets. In: *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*. vol. 10953, p. 109531Z. International Society for Optics and Photonics (2019)
5. Cohen, J.P., Morrison, P., Dao, L., Roth, K., Duong, T.Q., Ghassemi, M.: Covid-19 image data collection: Prospective predictions are the future. arXiv preprint arXiv:2006.11988 (2020)
6. Colleoni, E., Moccia, S., Du, X., De Momi, E., Stoyanov, D.: Deep learning based robotic tool detection and articulation estimation with spatio-temporal layers. *IEEE Robotics and Automation Letters* **4**(3), 2714–2721 (2019)
7. Haghani, A., Majdabadi, M.M., Ko, S.: Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning. arXiv preprint arXiv:2006.13807 (2020)
8. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 4700–4708 (2017)
9. Liu, H., Wang, L., Nan, Y., Jin, F., Wang, Q., Pu, J.: Sdfn: Segmentation-based deep fusion network for thoracic disease classification in chest x-ray images. *Computerized Medical Imaging and Graphics* **75**, 66–73 (2019)
10. Mahendran, A., Vedaldi, A.: Understanding deep image representations by inverting them. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 5188–5196 (2015)
11. Mangal, A., Kalia, S., Rajgopal, H., Rangarajan, K., Nambodiri, V., Banerjee, S., Arora, C.: Covidaid: Covid-19 detection using chest x-ray. arXiv preprint arXiv:2004.09803 (2020)
12. Marín, D., Aquino, A., Gegúndez-Arias, M.E., Bravo, J.M.: A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Transactions on medical imaging* **30**(1), 146–158 (2010)
13. McKeever, A.: Here’s what coronavirus does to the body. *National Geographic* (2020)
14. Moccia, S., Banali, R., Martini, C., Muscogiuri, G., Pontone, G., Pepi, M., Caiani, E.G.: Development and testing of a deep learning-based strategy for scar segmentation on cmr-lge images. *Magnetic Resonance Materials in Physics, Biology and Medicine* **32**(2), 187–195 (2019)

15. Organization, W.H., et al.: Health topics. coronavirus. Coronavirus: symptoms. World Health Organization, 2020a. Disponível em: https://www.who.int/healthtopics/coronavirus#tab=tab_3. Acesso em 7 (2020)
16. Öztürk, Ş., Özkaya, U., Barstuğan, M.: Classification of coronavirus (covid-19) from x-ray and ct images using shrunken features. *International Journal of Imaging Systems and Technology* (2020)
17. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R.: Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in Biology and Medicine* p. 103792 (2020)
18. Rosebrock, A.: Imagenet: Vggnet, resnet, inception, and xception with keras. *Mars* (2017)
19. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision*. pp. 618–626 (2017)
20. Spadea, M.F., Pileggi, G., Zaffino, P., Salome, P., Catana, C., Izquierdo-Garcia, D., Amato, F., Seco, J.: Deep convolution neural network (dcnn) multiplane approach to synthetic ct generation from mr images—application in brain proton therapy. *International Journal of Radiation Oncology* Biology* Physics* **105**(3), 495–503 (2019)
21. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 2818–2826 (2016)
22. Wang, L., Wong, A.: Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *arXiv preprint arXiv:2003.09871* (2020)
23. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: 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*. pp. 2097–2106 (2017)
24. Wu, Z., McGoogan, J.M.: Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention. *Jama* **323**(13), 1239–1242 (2020)
25. Zaffino, P., Pernelle, G., Mastmeyer, A., Mehrtash, A., Zhang, H., Kikinis, R., Kapur, T., Spadea, M.F.: Fully automatic catheter segmentation in mri with 3d convolutional neural networks: application to mri-guided gynecologic brachytherapy. *Physics in Medicine & Biology* **64**(16), 165008 (2019)
26. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: *European conference on computer vision*. pp. 818–833. Springer (2014)
27. Zhou, B., Sun, Y., Bau, D., Torralba, A.: Interpretable basis decomposition for visual explanation. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. pp. 119–134 (2018)
28. Zhou, J., Liao, X., Cao, J., Ling, G., Long, Q., et al.: Differential diagnosis between the coronavirus disease 2019 and streptococcus pneumoniae pneumonia by thin-slice ct features. *Clinical Imaging* (2020)
29. Zotin, A., Hamad, Y., Simonov, K., Kurako, M.: Lung boundary detection for chest x-ray images classification based on glcm and probabilistic neural networks. *Procedia Computer Science* **159**, 1439–1448 (2019)