Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 225 (2023) 227–236

Procedia Computer Science

www.elsevier.com/locate/procedia

27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

Deep feature analysis in a transfer learning approach for the automatic COVID-19 screening using chest X-ray images

Daniel I. Morís[a,b], Joaquim de Moura[a,b,*], Jorge Novo[a,b], Marcos Ortega[a,b]

[a] *Centro de Investigación CITIC, Universidade da Coruña, A Coruña, Spain*
[b] *VARPA Research Group, Instituto de Investigación Biomédica de A Coruña (INIBIC), Universidade da Coruña, A Coruña, Spain*

Abstract

COVID-19 is a challenging disease that was declared as global pandemic in March 2020. As the main impact of this disease is located in the pulmonary regions, chest X-ray devices are very useful to understand the severity of the disease on each patient. In order to reduce the risk of cross-contamination, the radiologists are recommended to use portable devices instead of fixed machinery, as these devices are easier to decontaminate. Moreover, the development of reliable and robust methodologies of computer-aided diagnosis systems is very relevant to reduce the workload that expert clinicians are experiencing in the current moment. In this work, we propose a comprehensive analysis of the deep features extracted from portable chest X-ray captures to perform a COVID-19 screening. We also study the optimal characterization of the problem with a lower dimensionality, contrasting the results of the feature selection methods that were chosen. Results demonstrated that the proposed approach is robust and reliable, obtaining a 90.43% of accuracy for the test set, using only 46.85% of the deep features in the context of poor quality and low detail X-ray images obtained from portable devices.

Keywords: computer-aided diagnosis; COVID-19; portable chest X-ray; deep learning; deep features

1. Introduction

The Coronavirus Disease 2019 (COVID-19) represents a great challenge for the health services worldwide [27], mainly caused due to the easy spread of the pathogen that causes this disease, the SARS-CoV-2. As consequence, the World Health Organization was forced to declare the COVID-19 as a global pandemic in 11th March 2020. To diagnose this disease, the gold-standard is the RT-PCR test [23]. However, this diagnostic tool is insufficient to know the extent and the severity of the COVID-19 on each patient. Despite the SARS-CoV-2 can harm several parts of the body, the

* Corresponding author. Tel.: +34 981167000 Ext. 6039
E-mail address: joaquim.demoura@udc.es

1877-0509
10.1016/j.procs.2023.10.007

(a) (b) (c)

Fig. 1: Examples of portable chest X-ray images of 3 different classes. (a) Normal class. (b) Pathological class. (c) COVID-19 class.

main affectation of this virus is located in the lung area. Therefore, chest X-ray imaging, a medical imaging modality that has been used to diagnose other common pulmonary pathologies as the pneumonia [7] or the tuberculosis [16] during the last decades, can be very useful to understand that state and the evolution of the disease for each case. To prevent the virus from spreading, radiologists are asked to prioritize the portable chest X-ray captures over any other image modality, because these devices are easier to decontaminate [14]. However, portable devices obtain captures with a lower level of detail, an aspect that makes it harder for clinicians to perform a reliable diagnosis.

Computer-aided diagnosis (CAD) methods are very useful to help the human experts in their diagnostic tasks. In particular, in the last years, the medical image analysis domain has benefited from the development of deep learning strategies, which are very powerful to deal with raw data. As COVID-19 diagnosis is a very relevant task, many contributions have addressed this problem. As reference, the work from Hussain *et al.* [11] extracts morphological and texture features from chest X-ray images. Then, the authors use those features to train classical machine learning algorithms. The contribution from Zhu *et al.* [30] uses deep network architectures to determine the severity of COVID-19 using chest X-ray images. In the work of De Moura *et al.*, 2022 [5], the authors prove the capabilities of several representative state-of-the-art deep learning models to solve the problem of COVID-19 screening using 3 different publicly available datasets. On its hand, the work from Wang *et al.* [28] proposes the use of a custom deep network tailored to the task of COVID-19 detection in chest X-ray images known as COVID-Net. Moreover, Keidar *et al.* [12] makes a comparison among several different deep learning models for COVID-19 detection in chest X-ray images. Apostolopoulos *et al.* [1] developed a methodology that extracts potentially representative COVID-19 biomarkers in chest X-ray images for a further classification. The proposal of Turkoglu [25] consists in the development of a methodology that performs a deep feature extraction using several parts of pre-trained networks to achieve a set of combined features. These combined features are then fed to a Support Vector Machine algorithm to distinguish between 3 classes: COVID-19, Normal and Pneumonia. Finally, it is worth to mention the contribution of Ullah *et al.* [26], that proposes a multi-task method to detect COVID-19 on chest X-ray images that is supported by other more well-known tasks (such as detection of pneumonia, pleural effusion or lung opacities) following a semi-supervised learning scheme that uses an adversarial autoencoder.

Specifically, given the relevance of portable chest X-ray imaging for the diagnosis of this relevant pandemic pathology, several computational approaches have been recently proposed. As reference, De Moura *et al.*, 2020 [3] developed a methodology focused on COVID-19 screening using portable chest X-ray images. Particularly, the authors used a dataset that was divided in 3 different classes: Normal, Pathological and COVID-19. Additionally, Morís *et al.*, 2021a [18] proposed data augmentation approaches using cycle-consistent adversarial networks for improving COVID-19 screening in portable chest X-ray images. In another proposal, Morís *et al.*, 2021b [17] contributed with a comprehensive analysis of the screening of COVID-19 approaches in chest X-ray images from portable devices, considering 6 representative state-of-the-art deep network architectures. Some examples of captures provided by this kind of portable devices can be seen in Fig. 1.

Despite the satisfactory results that were obtained in these works, most of them only partially address this relevant problem of global interest. Moreover, many of these works have some important limitations which are present in Convolutional Neural Network (CNN) architectures. Particularly, this kind of neural networks require high amounts of data and a great computational capability to deal with the back-propagation step which is necessary to train deep models. In addition, as the considered CNN models have layers with a notably high dimensionality, the works lack of

a detailed behavioral analysis of the extracted deep features. This kind of analysis is very important to know which is the set of features that obtains the optimum separability among classes reducing, at the same time, the original dimensionality of the problem. Particularly, in the context of medical imaging, deep feature extraction has been explored as an alternative to end-to-end approaches in domains such as chest X-ray [6], Computerized Tomography (CT) [20] or Optical Coherence Tomography (OCT) [4], among many others.

Given this important gap in the literature, in this work, we propose a comprehensive deep features analysis in a transfer learning-based approach for the automatic COVID-19 screening using portable chest X-ray images. For this purpose, we design a fully automatic methodology to distinguish X-ray images between 2 different classes: NON-COVID-19 and COVID-19. In particular, the proposed approach performs a deep feature extraction using several different pre-trained CNN models (VGG-16, AlexNet, Inception-V3 and ResNet-18), where the deep features are extracted from different fully-connected layers of these models. Apart from that, we also perform an exhaustive analysis of the optimal feature selection strategy that best characterizes the problem with a lower dimensionality and facilitates the learning process. This is supported by the removal of redundant and unnecessary features. To this aim, 3 different feature selection methods are selected: ReliefF, CFS and LLCFS. Afterward, we train and test a conventional machine learning model (a Support Vector Machine, abbreviated as SVM) to perform the classification using the previously selected deep feature set, considering a transfer learning strategy that reduces the computing capability requirements and simplifies the complexity of the original problem. In particular, the results demonstrate a competitive performance, with a 0.9043 of accuracy using only 46.86% of the whole set of features. This is a desirable characteristic for clinical settings, given the usual low capacity of the computational resources that are available in these environments, as they usually lack of advanced Graphics Processing Units (GPUs) to support the load of complex deep learning pipelines. Despite the existence of many works that address the problem of COVID-19 screening, to the best of our knowledge, this is the only work that addresses such an exhaustive analysis of the deep features extracted from an automatic screening model to detect COVID-19 using portable chest X-ray images.

2. Methodology

The methodology herein proposed, which is schematically described in Fig. 2, is divided in 3 main steps: "Deep Features Extraction", "Selection of Useful Features" and "Classification". Each of these steps is discussed in depth below.

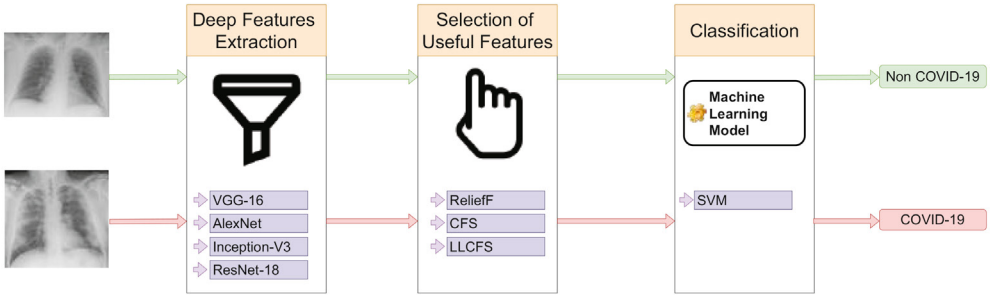


Fig. 2: Main overview of the proposed methodology, composed of 3 different steps: (I) deep features extraction, (II) selection of useful features and (III) classification.

2.1. Deep Features Extraction

Convolutional Neural Networks (CNN) have a particular architectural design that is suitable to reduce the dimensionality of a specific problem, extracting deep features from images. Then, this set of features can be used to perform a classification process [8]. In this work, we consider 4 different and representative deep network architectures pre-trained in the ImageNet dataset for the deep feature extraction: VGG-16, AlexNet, Inception-V3 and ResNet-18. Moreover, we also choose several different fully-connected layers of each architecture, to make a more detailed analysis and to obtain a greater performance of the method.

1. *VGG-16 model* [21]. This architecture is composed of 16 weighted layers. These layers can be divided in 13 convolutional layers and 3 fully-connected layers. The fully-connected layers are named as FC6, FC7 and FC8, having a dimensionality of 4096, 4096 and 1000 features, respectively.
2. *AlexNet model* [15]. This architecture is divided in different layers, namely, convolutional and pooling layers, rectified linear unit (ReLU) layers (that add non-linearity) and 3 fully-connected layers which are denoted as FC6, FC7, FC8, having 4096, 4096 and 1000 features, respectively.
3. *Inception-V3 model* [22]. This deep model, similarly as in the previous case, is composed of several cascaded convolutional and pooling layers, and a final fully-connected layer with 1000 features which is denoted as FC1000.
4. *ResNet-18 model* [10]. This kind of deep network architecture was proposed to overcome the vanishing gradient problem [24] adding identity shortcut connections in certain parts of the architecture, skipping one or more layers. The particular implementation herein considered has a depth of 18 layers.

2.2. Selection of Useful Features

One of the most important steps while using the deep features extracted from pre-trained models is the selection of the optimal subset of features that maximizes the separability between classes [2]. This deep feature selection step is very advantageous due to 2 main reasons. First, it reduces the size of the deep features set, thus minimizing the computational requirements and the time spent in the learning process. On the other hand, the selection of the optimal features removes irrelevant information that could significantly reduce its performance. In this way, to do the deep feature selection, 3 different algorithms were considered, based on the criteria of a previous related work that performed deep feature extraction with OCT images [4]: ReliefF, CFS (Correlation based Feature Selection) and LLCFS (Learning-based Clustering Feature Selection). A detailed explanation of each feature selection algorithm can be seen below:

1. *ReliefF* is a commonly used feature selection algorithm and is an extension of the Relief strategy [13]. ReliefF computes a score for each feature. Once the values are computed for the whole set of features, they can be ranked to select the top features that are more likely to be useful for the proposed problem. In order to do so, the algorithm computes the difference between the probabilities. In this way, feature scores are computed as specified in Eq. 1:

$$r(A) = P(A = Z|C) - P(A = Z|D) \tag{1}$$

   where Z refers to a specific value of the attribute A, C is the nearest neighbor to the current value that belongs to a different class and D is the nearest neighbor that belongs to the same class.
2. *CFS* algorithm [9] (Correlation-based feature selection) is based on the idea that the features with the greatest discriminatory capability are the most representative to predict the class, besides avoiding features that have correlations among them. With this definition, the expression of the CFS strategy is defined as stated in Eq. 2:

$$r_{zc} = \frac{k\overline{r_{zi}}}{\sqrt{k + k(k-1) * \overline{r_{ii}}}} \tag{2}$$

   where k denotes the amount of features, $\overline{r_{zi}}$ refers to the average correlation that exists between the ensemble of features and the sample label while the value $\overline{r_{ii}}$ express the average value of the inter-correlation that exists between features.
3. *LLCFS* is a feature selection algorithm based on the Local Learning-Based Clustering (LLC) [29] that uses the concept of kernel learning. The LLCFS method gives each feature (that can also be called as kernel in this context) a specific weight value, which is then included in the regularization of the LLC algorithm to reflect the importance of each feature in the problem. To do so, the considered function to perform the feature selection process is the local discriminant function that is shown in Eq. 3:

$$f_i^c(A) = A^{\mathrm{T}} w_i^c + b_i^c \tag{3}$$

where A denotes an arbitrary attribute, w_i^c denotes the regression coefficient given the cluster c and the i^{th} feature and b_i^c the same, but with the bias terms.

2.3. Classification

After the appropriate set of deep features is obtained, we train a suitable classification model, the Support Vector Machine (SVM) [19], following the same criteria as in [4]. In this context, the objective of the selected algorithm is to maximize a margin hyperplane to properly differentiate between 2 classes: COVID-19 and NON-COVID-19.

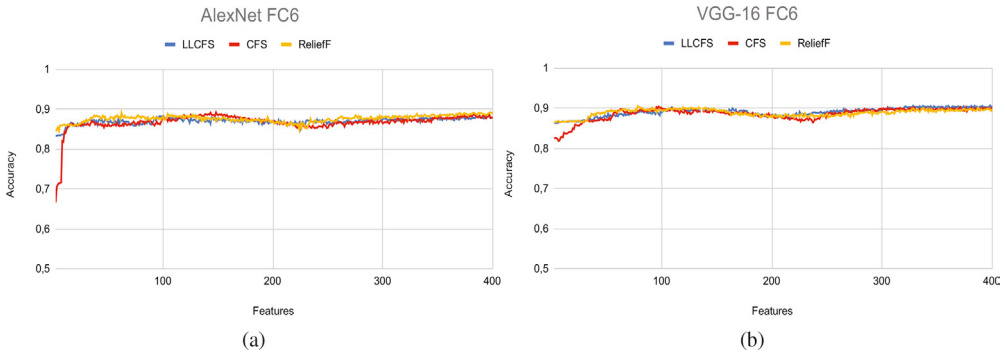
3. Results and Discussion

The developed methodology was validated using a dataset provided by the Complexo Hospitalario Universitario de A Coruña (CHUAC), that was specifically designed for the purposes of this work. The images were captured by 2 models of portable chest X-ray devices: Agfa dr100E and Optima Rx200. The set of images has a wide amount of different resolutions in the range between 712×742 and 1554×1910 pixels. Originally, the dataset is divided in 3 different classes: Normal, Pathological and COVID-19. The class Normal is composed of 717 samples from patients without evidences of being affected by pulmonary diseases that, however, could present pathological affectation in other regions. The class Pathological is composed of 717 samples obtained from patients that have evidences of pulmonary affectation from diseases different from COVID-19. Finally, the class COVID-19 is composed of 717 genuine COVID-19 cases. For the methodology validation purposes, the class Normal and the class Pathological are merged together to create a new class NON-COVID-19, while the class COVID-19 remains the same. Therefore, we train with a total of 1434 images for the NON-COVID-19 class and 717 images for the COVID-19 class, that keeps a proportion of $\frac{2}{3}$ and $\frac{1}{3}$ between classes, respectively. The input dataset is randomly divided in 2 different sets, with a 70% of the samples for training and the remaining 30% of samples for testing. Moreover, for the classification, we used a 10-fold cross-validation. In the same way, it is important to remark that the process is repeated 5 times and, in order to understand the global behavior of the models, the mean accuracy and the corresponding standard deviation are computed. One important aspect of the experimental procedure is that all the experiments were performed using different architectural settings, particularly, AlexNet, VGG-16, Inception-V3 and ResNet-18. In the same way, we use 3 different feature selection algorithms, in particular, ReliefF, CFS and LLCFS. Moreover, other element that we change is the fully-connected layer that is selected: FC6, FC7 or FC8/FC1000.

In this validation process, firstly, we use the feature selection algorithms to achieve the optimal subset of deep features. In Fig. 3, 4 and 5 we depict the performance evolution as the number of deep features used to train the models increases. There, the evolution is only shown for 400 features, as there is no further improvement with a greater number of deep features. This aspect can be seen notably in Fig. 6, where the performance of all the models is depicted. In fact, from 400 features onwards, the performance starts to experience a slight drop. However, in that part of the training process, the accuracy values are very stable. In the same way, the performance of the models is very similar and satisfactory in all the cases.

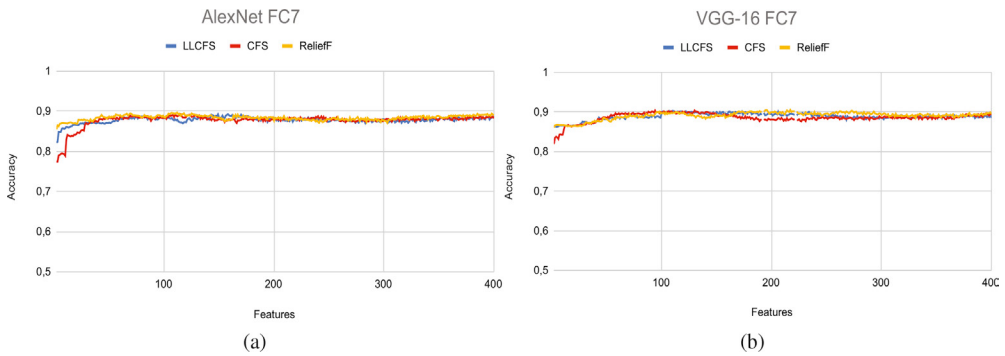
On the other hand, the Tables 1, 2 and 3 show the optimal feature subset size for each classification model configuration that have been obtained for the training stage. The overall summarized idea of the results is that all the configurations obtain a satisfactory performance. Particularly, the best performance using the features that were extracted from the FC6 layer was obtained with the VGG-16 model and the LLCFS feature selection algorithm with a 90.66% of mean accuracy, as it is shown in Table 1. In the case of the FC7 layer, the best performance is obtained using the VGG-16 architecture and the ReliefF algorithm with a mean accuracy of 90.62% as it is depicted in Table 2. Finally, for the case of the features that were selected from the FC8/FC1000 layer, the best performance is once again obtained using the same configuration as in the first scenario (VGG-16 and LLCFS feature selection algorithm) with a mean accuracy of 90.41% as can be seen in Table 3. With regard to the size of the optimal subsets, it was of 2129 features in the case of the FC6 layer, 3584 features in the case of the FC7 layer and 655 features in the case of the FC8/FC1000 layer.

In addition, the performance in test is shown in Tables 4, 5 and 6. There, it can be seen that the results depict a similar scenario as in the case of the training stage. All models are able to obtain satisfactory results, discriminating between NON-COVID-19 and COVID-19 using the optimal subset of deep features. In this case, the best performance



(a) (b)

Fig. 3: Accuracy evolution with respect to the number of used deep features that were obtained from the fully-connected layer FC6. (a) AlexNet model. (b) VGG-16 model.



(a) (b)

Fig. 4: Accuracy evolution with respect to the number of used deep features that were obtained from the fully-connected layer FC7. (a) AlexNet model. (b) VGG-16 model.

Table 1: Accuracy results achieved during the training stage by each deep model considering the optimal subset of features extracted by the fully-connected layer denoted as FC6.

Deep Model	FC6					
	Feature selector	*Nº of features*	*Min. Accuracy*	*Max. Accuracy*	*Mean Accuracy*	*Std. Accuracy*
AlexNet	CFS	1919	0.8845	0.8971	0.8919	0.0048
AlexNet	ReliefF	3666	0.8805	0.9011	0.8912	0.0075
AlexNet	LLCFS	2529	0.8712	0.8991	0.8831	0.0101
VGG-16	CFS	2390	0.8958	0.9137	0.9031	0.0069
VGG-16	ReliefF	927	0.8991	0.9117	0.9053	0.0049
VGG-16	LLCFS	2129	0.8997	0.9163	0.9066	0.0071

Table 2: Accuracy results obtained during the training stage achieved by each deep model considering the optimal subset of features extracted by the fully-connected layer FC7.

Deep Model	FC7					
	Feature selector	*Nº of features*	*Min. Accuracy*	*Max. Accuracy*	*Mean Accuracy*	*Std. Accuracy*
AlexNet	CFS	3985	0.8838	0.9024	0.8900	0.0075
AlexNet	ReliefF	501	0.8858	0.9037	0.8926	0.0069
AlexNet	LLCFS	846	0.8858	0.9037	0.8919	0.0072
VGG-16	CFS	562	0.8851	0.9077	0.8958	0.0085
VGG-16	ReliefF	3584	0.9011	0.9117	0.9062	0.0038
VGG-16	LLCFS	3410	0.8971	0.9077	0.9008	0.0053

obtained for the features extracted from the FC6 layer is achieved with the VGG-16 model and the LLCFS feature selection algorithm, with an accuracy of 90.43%. For the case of the features extracted from the FC7 layer, the best performance was achieved by the VGG-16 with the ReliefF algorithm as feature selector, with an accuracy of 89.44%.

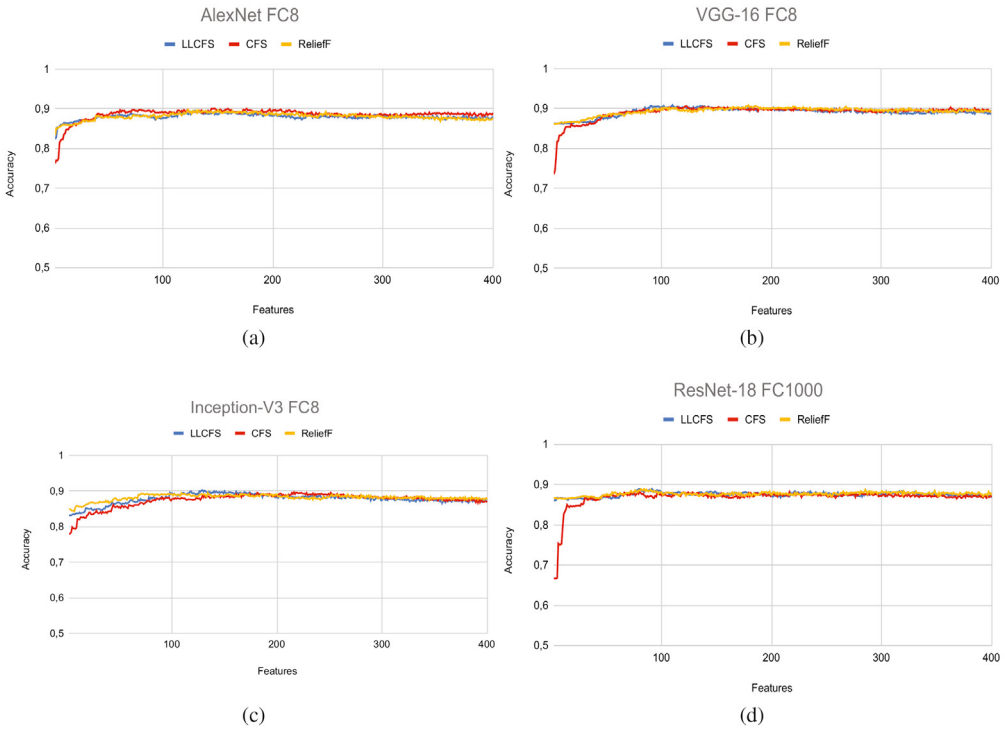


Fig. 5: Accuracy evolution with respect to the number of used deep features that were obtained from the fully-connected layer FC8/FC1000. (a) AlexNet model. (b) VGG-16 model. (c) Inception-V3 model. (d) ResNet-18 model.

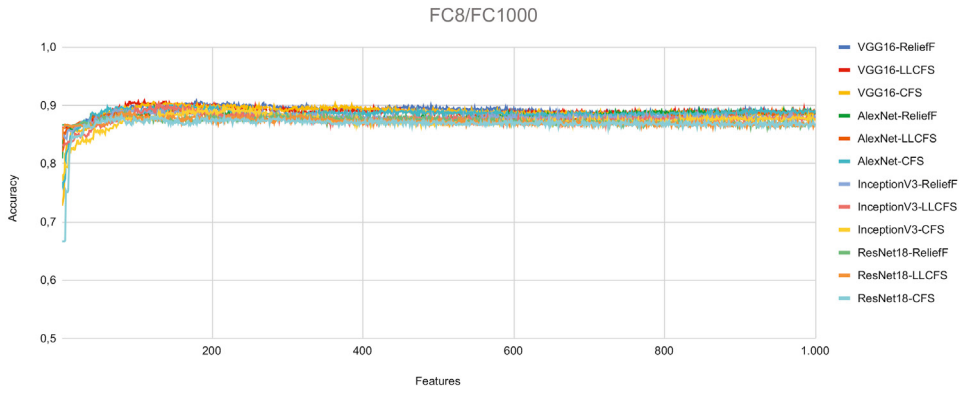


Fig. 6: Evolution of the accuracy values training with the whole subsets of deep features that were obtained from the used CNN networks, extracting the deep features from the fully-connected layers FC8/FC1000.

Finally, in the case of the FC8/FC1000 layer, the best accuracy value is achieved by 2 different configurations: using the AlexNet with the CFS feature selection algorithm and using the VGG-16 with the CFS algorithm. In both cases, the obtained mean accuracy value was of 88.84%. With respect to the number of features used by each configuration, it is remarkable that only the 46.85% and the 12.23% of the deep features from the FC6 and the FC7 layers, respectively, were necessary. With regard to the 2 different configurations in the case of the FC8/FC1000 that achieve the best performance, the AlexNet with the CFS feature selection algorithm uses only 65.50% of the whole amount of features while the VGG-16 with the CFS algorithm uses the 17.40% of the deep features. As it can be seen, the

Table 3: Accuracy results achieved during the training stage by each deep model considering the optimal subset of features extracted by the fully-connected layer denoted as FC8 or FC1000.

Deep Model	FC8/FC1000					
	Feature selector	*Nº of features*	*Min. Accuracy*	*Max. Accuracy*	*Mean Accuracy*	*Std. Accuracy*
AlexNet	CFS	290	0.8931	0.9057	0.8988	0.0046
AlexNet	ReliefF	52	0.8765	0.9037	0.8871	0.0101
AlexNet	LLCFS	174	0.8858	0.9017	0.8904	0.0066
VGG-16	CFS	184	0.8964	0.9104	0.9015	0.0058
VGG-16	ReliefF	163	0.8931	0.9117	0.9020	0.0068
VGG-16	LLCFS	655	0.9004	0.9130	0.9041	0.0053
Inception-V3	CFS	158	0.8884	0.9050	0.8965	0.0060
Inception-V3	ReliefF	223	0.8838	0.9011	0.8915	0.0066
Inception-V3	LLCFS	819	0.8977	0.9084	0.9023	0.0041
ResNet-18	CFS	64	0.8705	0.8898	0.8786	0.0073
ResNet-18	ReliefF	573	0.8772	0.8997	0.8846	0.0089
ResNet-18	LLCFS	292	0.8712	0.8971	0.8813	0.0101

Table 4: Best accuracy values achieved in test using the different deep network architectures and the optimal subset of features extracted from the fully-connected FC6 layer.

Deep Model	FC6		
	Feature selector	*% of Features*	*Accuracy*
AlexNet	CFS	51.97%	0.8818
AlexNet	ReliefF	22.63%	0.8884
AlexNet	LLCFS	58.34%	0.8711
VGG-16	CFS	61.74%	0.9023
VGG-16	ReliefF	89.50%	0.8857
VGG-16	LLCFS	46.85%	0.9043

Table 5: Best accuracy values obtained in test using the considered deep network architectures and the subset of features extracted from the fully-connected FC7 layer.

Deep Model	FC7		
	Feature selector	*% of Features*	*Accuracy*
AlexNet	CFS	83.25%	0.8818
AlexNet	ReliefF	46.85%	0.8917
AlexNet	LLCFS	13.72%	0.8758
VGG-16	CFS	20.65%	0.8857
VGG-16	ReliefF	12.23%	0.8944
VGG-16	LLCFS	97.29%	0.8778

Table 6: Best accuracy obtained during the test stage using the different deep model architectures and the optimal subset of features extracted from the fully-connected FC8/FC1000 layer.

Deep Model	FC8/FC1000		
	Feature selector	*% of Features*	*Accuracy*
AlexNet	CFS	65.50%	0.8884
AlexNet	ReliefF	16.30%	0.8837
AlexNet	LLCFS	18.40%	0.8844
VGG-16	CFS	17.40%	0.8884
VGG-16	ReliefF	5.20%	0.8851
VGG-16	LLCFS	29.00%	0.8837
Inception-V3	CFS	81.90	0.8778
Inception-V3	ReliefF	22.30%	0.8764
Inception-V3	LLCFS	15.80%	0.8764
ResNet-18	CFS	29.20%	0.8698
ResNet-18	ReliefF	57.30%	0.8718
ResNet-18	LLCFS	6.40%	0.8731

best performance obtained using the features extracted from the FC6 layer outperforms the best performance obtained using both the FC7 and the FC8/FC1000 layers. This is due to the fact that the FC6 is a shallower layer, having a more global representation of the overall condition and more global characteristics of the samples. However, the performance drop is slightly noticeable. Thus, the results of the methodology proposed in this contribution prove to be robust and reliable, despite using a representation with a lower dimensionality. Therefore, this approach has proven to be efficient in terms of memory requirements.

The scientific community has been making significant efforts to address the problem of COVID-19 screening using portable chest X-ray images. To the best of our knowledge, this study represents the first comprehensive analysis of deep features extracted from multiple architectures for this purpose. It is important to note that portable chest X-ray images tend to have lower quality compared to traditional chest X-ray images, but they were widely used during the

COVID-19 pandemic due to the need for fast and accessible screening. Despite this challenge, our proposed method achieved competitive performance while utilizing only a fraction of the features, demonstrating the efficiency of the system and reducing memory requirements compared to end-to-end approaches. This aspect is desirable in clinical settings, that usually struggle to implement the end-to-end approaches in daily practice due to the lack of advanced computing resources. Furthermore, this approach has the potential for application in various other biomedical imaging modalities. It is important to note that there is no possibility to compare our proposed approach in fair conditions with other works of the state-of-the-art, given that the dataset herein used was specifically tailored for the study and is unavailable publically. In the same line, to the best of our knowledge, this is the only dataset that has been tailored to perform studies with portable chest X-ray images exclusively. Nevertheless, it is remarkable that the achieved results are in line with the performance of the state-of-the-art approaches.

4. Conclusions

In this work, we proposed an efficient fully automatic methodology to perform an exhaustive analysis of the deep features extracted from portable chest X-ray images of COVID-19 patients to distinguish between 2 classes: NON-COVID-19 and COVID-19. Therefore, the method herein proposed can provide a useful tool to reduce the workload that health care services are suffering due to the global pandemic of COVID-19. To do so, we have considered a pipeline that can be divided in 3 different steps. Firstly, we extract a set of deep features with several different remarkable deep network architectures: AlexNet, VGG-16, Inception-V3 and ResNet-18. These features are extracted from several fully-connected layers from each deep model. In particular, we use the FC6 and the FC7 from the AlexNet and the VGG-16 models as well as the FC8/FC1000 from the AlexNet, the VGG-16, the Inception-V3 and the ResNet-18. Secondly, we perform an exhaustive analysis of the set of extracted deep features, to determine the set of features that optimally separates between the 2 classes. To do so, we use 3 different and representative feature selection algorithms: the ReliefF, the LLCFS and the CFS. Finally, the different subsets of features are used to train an SVM model, that performs the image classification. Despite the complexity of the proposed scenario, the model is able to accurately separate the NON-COVID-19 and the COVID-19 cases, achieving the best global accuracy value of 90.43% in test. The best result was obtained with the deep features extracted from the VGG-16 pre-trained model and filtered with the LLCFS feature selection algorithm, using only the 46.85% of the whole amount of features retrieved from the FC6 layer. Therefore, the method proves its robustness dealing with the proposed scenario, providing a better understanding of the problem thanks to the analysis of the deep features. In addition, this work represents a powerful alternative to optimize the use of computational resources and to reduce the complexity of the original problem. This is a critical aspect in clinical environments, given that the available computational resources are usually of low capacity. As possible lines of future work, this analysis could be complemented evaluating the performance obtained with other machine learning classifiers.

Acknowledgements

This work was supported by Ministerio de Ciencia e Innovación, Government of Spain through the research project with [grant numbers RTI2018-095894-B-I00, PID2019-108435RB-I00, TED2021-131201B-I00, and PDC2022-133132-I00]; Consellería de Educación, Universidade, e Formación Profesional, Xunta de Galicia, Grupos de Referencia Competitiva, [grant number ED431C 2020/24], predoctoral grant [grant number ED481A 2021/196]; CITIC, Centro de Investigación de Galicia [grant number ED431G 2019/01], receives financial support from Consellería de Educación, Universidade e Formación Profesional, Xunta de Galicia, through the ERDF (80%) and Secretaría Xeral de Universidades (20%).

References

[1] Apostolopoulos, I.D., Aznaouridis, S.I., Tzani, M.A., 2020. 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 , 1–14.doi:10.48550/arXiv.2004.00338.

[2] Chandrashekar, G., Sahin, F., 2014. A survey on feature selection methods. Computers & Electrical Engineering 40, 16–28. doi:10.1016/j.compeleceng.2013.11.024. 40th-year commemorative issue.
[3] De Moura, J., García, L.R., Vidal, P.F.L., Cruz, M., López, L.A., Lopez, E.C., Novo, J., Ortega, M., 2020. Deep convolutional approaches for the analysis of covid-19 using chest x-ray images from portable devices. IEEE Access 8, 195594–195607. doi:10.1109/ACCESS.2020.3033762.
[4] De Moura, J., Novo, J., Ortega, M., 2019. Deep feature analysis in a transfer learning-based approach for the automatic identification of diabetic macular edema, in: 2019 International Joint Conference on Neural Networks (IJCNN), IEEE. doi:10.1109/ijcnn.2019.8852196.
[5] De Moura, J., Novo, J., Ortega, M., 2022. Fully automatic deep convolutional approaches for the analysis of covid-19 using chest x-ray images. Applied Soft Computing 115, 108190. doi:10.1016/j.asoc.2021.108190.
[6] Devnath, L., Luo, S., Summons, P., Wang, D., 2021. Automated detection of pneumoconiosis with multilevel deep features learned from chest x-ray radiographs. Computers in Biology and Medicine 129, 104125. doi:10.1016/j.compbiomed.2020.104125.
[7] Gennis, P., Gallagher, J., Falvo, C., Baker, S., Than, W., 1989. Clinical criteria for the detection of pneumonia in adults: guidelines for ordering chest roentgenograms in the emergency department. The Journal of emergency medicine 7, 263–268. doi:10.1016/0736-4679(89)90358-2.
[8] Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press. http://www.deeplearningbook.org.
[9] Hall, M.A., 1999. Correlation-based feature selection for machine learning. doi:10.1109/CESYS.2018.8723980.
[10] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778. doi:10.48550/arXiv.1512.03385.
[11] Hussain, L., Nguyen, T., Li, H., Abbasi, A.A., Lone, K.J., Zhao, Z., Zaib, M., Chen, A., Duong, T.Q., 2020. Machine-learning classification of texture features of portable chest x-ray accurately classifies covid-19 lung infection. BioMedical Engineering OnLine 19, 1–18. doi:10.1186/s12938-020-00831-x.
[12] Keidar, D., Yaron, D., Goldstein, E., Shachar, Y., Blass, A., Charbinsky, L., Aharony, I., Lifshitz, L., Lumelsky, D., Neeman, Z., et al., 2021. Covid-19 classification of x-ray images using deep neural networks. European radiology , 1–10.doi:10.1007/s00330-021-08050-1.
[13] Kira, K., Rendell, L.A., 1992. A practical approach to feature selection, in: Sleeman, D., Edwards, P. (Eds.), Machine Learning Proceedings 1992. Morgan Kaufmann, San Francisco (CA), pp. 249–256. doi:10.1016/B978-1-55860-247-2.50037-1.
[14] Kooraki, S., Hosseiny, M., Myers, L., Gholamrezanezhad, A., 2020. Coronavirus (covid-19) outbreak: what the department of radiology should know. Journal of the American college of radiology 17, 447–451. doi:10.1016/j.jacr.2020.02.008.
[15] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25, 1097–1105. doi:10.1145/3065386.
[16] Mert, A., Bilir, M., Tabak, F., Ozaras, R., Ozturk, R., Senturk, H., Aki, H., Seyhan, N., Karayel, T., Aktuglu, Y., 2001. Miliary tuberculosis: clinical manifestations, diagnosis and outcome in 38 adults. Respirology 6, 217–224. doi:10.1046/j.1440-1843.2001.00328.x.
[17] Morís, D.I., de Moura, J., Novo, J., Ortega, M., 2021a. Comprehensive analysis of the screening of COVID-19 approaches in chest x-ray images from portable devices, in: ESANN 2021 proceedings, Ciaco - i6doc.com. doi:10.14428/esann/2021.es2021-31.
[18] Morís, D.I., de Moura, J., Novo, J., Ortega, M., 2021b. Data augmentation approaches using cycle-consistent adversarial networks for improving covid-19 screening in portable chest x-ray images. Expert Systems with Applications 185, 115681. doi:10.1016/j.eswa.2021.115681.
[19] Noble, W.S., 2006. What is a support vector machine? Nature biotechnology 24, 1565–1567. doi:10.1038/nbt1206-1565.
[20] Samala, R.K., Chan, H.P., Hadjiiski, L.M., Richter, C., Zhou, C., Wei, J., 2019. Analysis of deep convolutional features for detection of lung nodules in computed tomography, in: Hahn, H.K., Mori, K. (Eds.), Medical Imaging 2019: Computer-Aided Diagnosis, SPIE. doi:10.1117/12.2512208.
[21] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv. doi:10.48550/arXiv.1409.1556.
[22] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826. doi:10.48550/arXiv.1512.00567.
[23] Tahamtan, A., Ardebili, A., 2020. Real-time rt-pcr in covid-19 detection: issues affecting the results. Expert review of molecular diagnostics 20, 453–454. doi:10.1080/14737159.2020.1757437.
[24] Tan, H.H., Lim, K.H., 2019. Vanishing gradient mitigation with deep learning neural network optimization, in: 2019 7th international conference on smart computing & communications (ICSCC), IEEE. pp. 1–4. doi:10.1109/ICSCC.2019.8843652.
[25] Turkoglu, M., 2021. Covidetectionet: Covid-19 diagnosis system based on x-ray images using features selected from pre-learned deep features ensemble. Applied Intelligence 51, 1213–1226. doi:10.1007/s10489-020-01888-w.
[26] Ullah, Z., Usman, M., Gwak, J., 2023. MTSS-AAE: Multi-task semi-supervised adversarial autoencoding for COVID-19 detection based on chest x-ray images. Expert Systems with Applications 216, 119475. doi:10.1016/j.eswa.2022.119475.
[27] Velavan, T.P., Meyer, C.G., 2020. The COVID-19 epidemic. Tropical Medicine & International Health 25, 278–280. doi:10.1111/tmi.13383.
[28] Wang, L., Lin, Z.Q., Wong, A., 2020. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports 10, 1–12. doi:10.1038/s41598-020-76550-z.
[29] Zeng, H., Cheung, Y.m., 2010. Feature selection and kernel learning for local learning-based clustering. IEEE transactions on pattern analysis and machine intelligence 33, 1532–1547. doi:10.1109/TPAMI.2010.215.
[30] Zhu, J., Shen, B., Abbasi, A., Hoshmand-Kochi, M., Li, H., Duong, T.Q., 2020. Deep transfer learning artificial intelligence accurately stages covid-19 lung disease severity on portable chest radiographs. PloS one 15, e0236621. doi:10.1371/journal.pone.0236621.