Digital Object Identifier

COVID-19 and pneumonia detection from chest X-ray images with Deep Learning

VITÓRIA CRUZ[1], JOÃO J. ALEGRIA[2]
[1]Departamento de Eletrónica, Telecomunicações e Informática, Universidade de Aveiro (nmec: 89189; e-mail: vicruz99@ua.pt)
[2]Departamento de Eletrónica, Telecomunicações e Informática, Universidade de Aveiro (nmec: 89243; e-mail: joaoalegria31@ua.pt)

ABSTRACT The goal of this work was to develop a deep learning model to classify X-Ray images of healthy patients, patients with Covid-19 and patients with pneumonia. This task can be hard due to the fact that if the disease is caught in an early stage, it's unlikely that anything substantial will appear on the X-Ray. We have used a dataset with 4725 samples, whose classes were balanced.
The models proposed in the literature to solve this problem fit one of the following groups: create a CNN from scratch, do transfer learning for feature extraction or do transfer learning with fine-tuning. All models that were reviewed registered accuracies above 90%.
We have implemented two models, one based on transfer learning for feature extraction and another based on transfer learning with fine tuning. In the first model, five pre-trained CNNs were used to extract features from the images in our dataset. Then, after feature reduction was applied, several linear classifiers were trained with these features. The results were disapointing. We couldn't determine the source of the problem, as it was shown that the features contained relevant information about the data.
In the second model the last layer of five pre-trained CNNs was substituted by a simpler layer containing just three neurons. An ensemble classifier based on hard voting was obtained from the CNNs. The results achieved were satisfactory. The CNN RestNet and the CNN DenseNet had the best performances. The accuracy of the modified ResNet, which was the model with the best performance, was 87.1%. This value is inferior to most values present in literature, but is is still very satisfactory.

I. INTRODUCTION

Coronaviruses are a family of viruses that can cause common colds, flu, fever, and other illnesses in humans. One of the most severe variants of this type of virus is SARS-Cov-2, also known as Covid-19. In December 2019, health institutions of Wuhan city in the Hubei province of China reported several unknown pneumonia cases [1]. It was then confirmed on January 7, 2020, that the viral pneumonia cases were caused by a novel coronavirus, which then spread around the world. In Portugal, about 3.24 million people have become infected with this virus, and 20 973 have died. Globally, these numbers rise to 434 million infections and 6 million deaths.

The most common test technique currently used for COVID-19 diagnosis is a real-time reverse transcription-polymerase chain reaction (RT-PCR). However, in the beginning of the pandemic, due to the low sensitivity of these tests and due to its frequent inavailability, doctors in Chinese clinics were encouraged to make a diagnosis only based on clinical and chest CT results [2], [3]. It has been shown that radiologic images of Covid-19 patients contain useful information for diagnostics of the disease and can serve as a viable alternative diagnosis method. The most common findings in X-rays seem to be ground glass opacities, strides or lines in the lungs. The use of X-Rays is thus an alternative to the RT-PRC tests, which are costly and whose results can come slowly.

There are various types of pneumonia, such as bacterial pneumonia and viral pneumonia, which can damage the lungs in different ways. Thus, it may be possible to distinguish X-Rays of patients with pneumonia and of patients with Covid-19, as different diseases damage the lungs in different ways. In Figure 1 we present X-Rays of patients in a healthy condition and with pneumonia, respectively, where an opaque region in the lugs is very visible, and indicative of pneumonia.

Due to the limited number of radiologists, it is difficult to provide every hospital with an expert clinician. This, among other reasons, has made automated detection systems based on AI techniques especially attractive, as they can provide a simple, accurate and fast way to aid doctors in their diagnosis of either Covid-19 and pneumonia. Additionally,

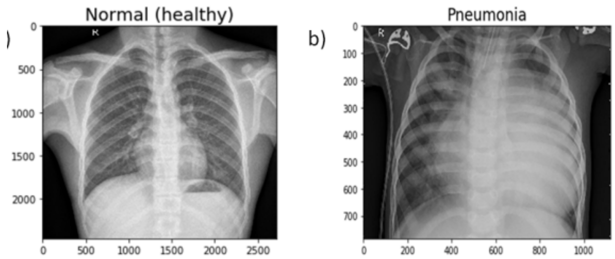

FIGURE 1: X-ray from an healthy patient (left) and with pneumonia (right)

one of the biggest disadvantages of using chest radiography for the detection of Covid-19, is its inability in detecting the disease in its early stages. However, AI models, and more specifically deep learning models, may notice details difficulty perceptible to the human eye and help solve this problem.

In the literature, several approaches have been proposed to solve this problem, most of which consist in using deep learning for feature extraction or fine tuning, or building a CNN from scratch. One of the thing that hinders the development of models for this particular application is the lack of a big dataset. Most models in literature have been trained only with a few thousand images.

In this work, we intend to develop one or more models that are able to classify a dataset of images of X-Rays of healthy people, people with pneumonia and people with Covid-19. Some of our models are adaptions of models that we have encountered in the literature, but which have only been applied to a binary classification case, such as Healthy vs Pneumonia or Healthy vs Covid-19. The models we intend to train are both based on transfer learning. In one of the models we will perform fine tuning of several pre-trained CNNs and create an ensemble classifier out of all the resulting CNNs. In the other model we will perform feature extraction by also using several pre-trained CNNs and then use a feature reduction method (mRMR). The resulting features will then be feed to various linear classifiers. The model with the highest performance will be the chosen one.

II. STATE OF THE ART

Most approaches presented in literature that aim to solve a classification problem based on the analysis of medical images resort to the use of deep learning, as this is a particularly suitable method when extracting features from raw data is a difficult talks, as is the case when images are used. Thus, it is expected that the majority of models used to classify X-Ray images are deep learning models.

The classification of images in the groups of Healthy, Pneumonia and Covid-19 is very specific problem. To cover a wider range of works, we have also analysed studies when the problem was just a binary classification problem, such as Healthy vs Pneumonia or Healthy vs Covid-19.

In the article "A Deep Feature Learning Model for Pneumonia Detection Applying a Combination of mRMR Feature Selection and Machine Learning Models" [4] the authors intended to solve a binary classification problem using a total of 5,849 images of healthy patients (1,583 images) and pneumonia patients (4,266 images). Three pre-trained CNN models, AlexNet, VGG16 and VGG19, trained on the ImageNet dataset, were used to perform feature extraction. Each pre-trained model contributed with 1000 features, which were the outputs of the Fully Connected Layers (FCL) of each model. Then, feature selection was performed using the mRMR method, more specifically the MID and MIC variants [5], and the best 100 features of each model were chosen. The resulting 300 features were then fed to different classifiers (Logistic Regression, SVM, Linear Discriminant). The Linear Discriminant method had the best performance, with an accuracy of 99.41%. The authors concluded that the features provided by the deep learning models were very informative, thus allowing the construction of a very robust model. Additionally, mRMR was also proved useful in the reduction of number of features. Some of the things we considered most attractive about this method were: the linear classifiers are light and easy train; because the classifiers are separated from the part of the model that performs feature extraction, there is a lot of flexibility in terms of the classifiers that can be used; once again, because the feature extraction is separated from the classifiers, there is a lot of flexibility in terms of which features can be feed into the classifiers.

In the article "A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images" [6] the authors intended to solve a binary classification problem using a total of 5,232 images of healthy patients (1,346 images) and pneumonia patients (3,886 images). In this dataset, the pneumonia images were divided by different types of pneumonia, which could make the classification task more difficult. In this work the authors used transfer learning and substituted the FCL layer of five different pre-trained CNN (AlexNet, DenseNet121, ResNet18, InceptionV3, GoogleLeNet), which had about 1000 neurons each, by a FCL of 2 neurons. Then, this layer was trained. The altered CNN with the best performance was the ResNet18. All CNNs were then gathered in an ensemble classifier with hard voting, which had the best performance among all the models, registering an accuracy of 96.4%. Although inferior to the result obtained in the previous paper, this is still a very high result. Within this paper there are also mentions to related works, and the authors conclude that the obtained results are superior to the results of many other approaches used to solve similar problems.

In the article "Efficient Pneumonia Detection in Chest Xray Images Using Deep Transfer Learning" [7] the authors intended once again to solve a binary classification problem using a total of 5836 images of healthy patients (1,583 images) and pneumonia patients (4,283 images). The approach used in this article is very similar to the one used in previous one. Five pre-trained CNNs (ResNet18, Xception, InceptionV3, DenseNet121, and MobileNetV3) were fine tuned and their predictions were combined in an ensemble classifier, with the difference that instead of being based in hard voting, this was a weighted classifier. Once again the ensemble classifier achieved the highest performance, with an accuracy of 98.43%, which is higher than the one obtained in the previous paper. From the individual models, the CNNs with the best performance was DenseNet121.

In the article "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare" [8] the authors intended once again to solve a binary classification problem using a total of 5836 images of healthy patients and pneumonia patients. The goal of this paper was to train a CNN from scratch to show that is possible to obtain a model to classify medical images that doesn't rely on transfer learning. To compensate for the low number of examples, image augmentation was used in this dataset. In total, the CNN had about 6.7 million parameters to tune, which is a very high number when compared to the number of available examples. The accuracy obtained with validation was 93.73%. The authors clearly stated that no testing set was used, which means this result may not describe accurately the real performance of the model.

In the article "X-ray and CT-scan-based automated detection and classification of covid-19 using convolutional neural networks (CNN)" [15] the authors intended to solve a multi-class classification problem using a total of 11,095 images of healthy patients, covid-19 patients and pneumonia patients. They have built a CNN from scratch and achieved an accuracy of 98.28%. The details of the CNN weren't disclosed, so we couldn't find an estimate for the number of parameters used. However, contrary to what was done in the previous paper, these authors have evaluated the performance of the model in a testing set. This dataset didn't just used images of X-Ray and used, as well, imagens of CTs, which is why the dataset used in this number was considerably bigger than the one used in other papers.

In the article "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images" [9], the authors intended to solve a multi-class classification problem using a total of 1,251 images of healthy patients, covid-19 patients and pneumonia patients. Due to the low number of examples, the authors fine-tuned the pre-trained Xception and achieved an accuracy of 90%.

In the article "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images" [10] the authors intended to solve a multi-class classification problem using a total of 4575 images of healthy patients (3050) and covid-19 patients (1525). In this paper a CNN, which was used as a feature extractor, was combined with a LTSM, which was used a classifier. This model achieved an accuracy of 99.4%, which is the best performance that we have encountered for a multi-class classifier.

In the following table we include a list of the methods previously presented and of their accuracy. Additionally, we've also included a few more models, so it's more clear what the average performance of the multiclass models is.

Type of Model	Type of Classification	Number of Examples	Accuracy
Transfer Learning (Feature Extraction)	Binary	5849	99.4%
Transfer Learning (Fine Tuning)	Binary	5232	96.4%
Transfer Learning (Fine Tuning)	Binary	5836	98.43%
CNN	Binary	5836	93.73%
CNN	Multiclass	11095	98.28%
Transfer Learning (Fine Tuning)	Multiclass	1251	90%
CNN + LTSM	Multiclass	4575	99.4%
Transfer Learning (Fine-Tuning) [1]	Multiclass	---	93.5%
CNN [2]	Multiclass	---	92.4%
CNN [3]	Multiclass	---	90.0%

TABLE 1: list of used methods and their correspondent accuracies

As expected, the performance of the binary models is overall better than the performance of the multiclass models.

There seems to be an adequate distribution between transfer learning based models and models trained from scratch, with no clear indication that any of these approaches produces better results than the other.

Although the generality of results is very good, one of the things that the authors indicate that is preventing better results is the lack of data, as most datasets contained only a limited amount of examples.

III. DATA DESCRIPTION AND VISUALISATION

The goal of this work is to assign X-Ray images to the following classes: healthy, pneumonia, covid-19.

The dataset [14], which was provided by Mendeley, contains chest X-ray posteroanterior (PA) RGB images of people with Covid-19, Pneumonia or with nothing. Each class has 1525 images. In total there are 4575 examples. About 900 images from the covid-19 class were obtained by image augmentation, which was done by the providers of the dataset to balance the classes. This means that there is less variety of covid-19 images, which may cause the models to not perform as well in this class.

Our dataset was divided in training (60% - 2745 images), validation (20% - 915 images) and testing (20% - 915 images). Due to the fact that it will be necessary to train neural networks, we couldn't afford to use cross-validation and join the training and validation subsets.

The number of examples in our dataset is inferior to the number of examples used in most models in the previous section. In Figure 2 we present some of the images present in the dataset. Some of the covid images are rotated due to the fact that they were obtained with image augmentation.

In most images of pneumonia and in some images of covid, it is clear that there is opacity in the lungs, indicating the presence of these diseases. In case of the images of pneumonia,

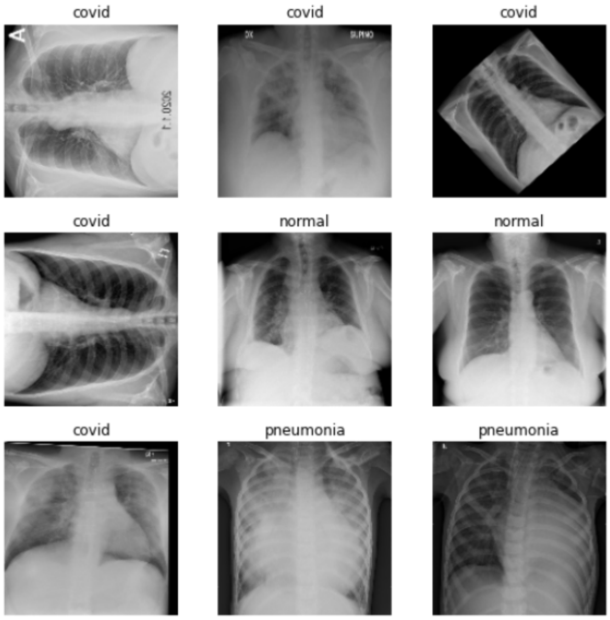

FIGURE 2: images from the dataset representing the different classes

this opacity seems to be dispersed through the lung. In case of the images of covid, this opacity seems to be localised in more specific regions. However, there also some images of covid that are fairly similar to the ones of healthy people, which probably correspond to cases where the disease is still in an early state. This can make it difficult to distinguish healthy people from some people with pneumonia or covid.

In some images there seems to be a label present. In the following figure we present an example, in which the label “L” is visible on the top of the figure.

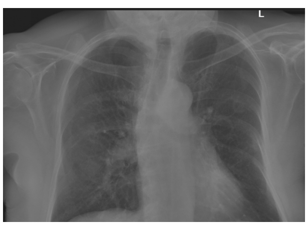

FIGURE 3: X-Ray image with a visible L on top of the figure

After a high level analysis of the dataset, we noticed this label was present in several images, along with other labels. If the distribution of these labels is similar across the different classes, this shouldn’t be a problem. However, if this is not the case, the models may inadvertently “learn” that those labels are indicative of classes, instead of “learning” to recognise signs of the disease in the image, which may make them not generalise well. Because the data has over four thousand images, we couldn’t confirm that this distribution is equal among classes, but after a high level analysis, it appeared to be.

IV. THE MODELS

In the State of the Art section, we analysed the approaches followed by several different studies to solve the problem of binary classification and multiclass detection of covid and/or pneumonia images. Due to the fact that this problem concerns the analysis of images, all the studies we came across used deep learning models. In general, the models presented in the studies were divided into the following sets: CNNs built from scratch, transfer learning with fine tuning and transfer learning with feature extraction.

In our case, building a model with a CNN from scratch doesn’t seem like a viable option, since the number of images we have is relatively small and some of them are the result of image augmentation, which means they don’t contain "new information". The network we would build could never be too complex, as we would run the risk of overfitting. However, the network’s lack of complexity would mean that it would likely not be able to solve the problem at hand. Thus, the use of transfer learning seems to be more suitable for our case.

There are two types of approaches that can be followed with transfer learning: transfer learning for feature extraction or transfer learning with fine-tuning. With transfer learning for feature extraction, the outputs of the pre-trained network are treated as features. These "outputs" do not have to be obtained at the last layer of the network, they can be obtained at any point. Classifiers are then trained with these features. With fine-tuned transfer learning, the last layer or layers of the network are usually replaced by simpler layers, which then have to be trained.

We decided to develop two models based on transfer learning. The models were inspired in models proposed in two articles presented in the State of the Art section. In these articles the models were used only for binary classification. Our aim is to extend these models to the case of multi-class classification.

A. TRANSFER LEARNING WITH FEATURE EXTRACTION

For transfer learning for feature extraction, we decided to follow an approach similar to the one proposed in the article [4], where the accuracy obtained for binary classification was very high 99.4%. In this work, the outputs of three pre-trained networks are considered features. For each network, 1000 features are obtained, as each pre-trained network has 1000 classes. The feature reduction method mRMR [5] is applied to each feature set, and 100 features for each model are selected. As three models were used, 300 features are obtained in the end. Several linear classifiers are then trained based on these features. Because the feature extraction is done independently of the classifiers, this approach offers a lot of flexibility in terms of the classifiers that will be used. Additionally, linear classifiers are very light and fast to train.

We have implemented a slightly different version of this approach.

The pre-trained models that we used to do feature extraction were: MobileNetV2, VGG16, VGG19, DenseNet121, ResNet50V2. We tried to select a wide range of models, from the ones that were made available by Keras, and which would probably have a good performance on the analysis of X-Ray images of Covid and Pneumonia [16].

We then passed the data from our dataset through these networks and saved the features present at the output of the last layer of the network. As all these networks had about 1000 classes, we got 5000 features. Before passing the images through the models, the images were scaled to the size 224x224, as this was the size accepted by the models, and were pre-processed according to the requirements of each model. We then applied a feature reduction method, mRMR. Succinctly, this method selects the features that have the maximum relevance with respect to the target variable and minimum redundancy (i.e. correlation) with respect to the features have been selected at previous iterations. More specifically, we have applied the FCD variant of this method, as it has been shown that this variant "has an outstanding performance in computation and robust accuracy for different down-stream classification methods" [5]. The user only needs to specify the number of features it desires before applying this method. Instead of choosing the best 100 features given by each model, we decided to select the best 10, 100, 200, 300, 400 and 500 features from the whole group of features, to see how the accuracy of the classifiers was affected by the number of features. After this, several different linear classifiers were trained (Logistic Regression, SVM, Linear Discriminant, Quadratic Discriminant, K-Nearest) and a search for their best hyper-parameters was performed. The hyper-parameters that were tested can be found in Table 2.

Models	Grid of Parameters
Logistic Regression	C = [0.001, 0.01, 1, 10, 100] Penalty = [L1, L2]
SVM	C = [0.01, 0.1, 1, 10, 100] gamma = [Auto] kernel = [RBF, Linear, Sigmoid]
Linear Discriminant	Tolerância = [None, 10^-12, 10^-10, 10^-6] Shrinkage = [None, Auto] Solver = [SVD, LSQR]
Quadratic Discriminant	Parâmetro de Regularização = [0.6, 0.7, 0.8, 0.9, 1.0]
K-Nearest	Nº of Neighbours = [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] weights = [uniform, distance]

* All the sklearn parameters presented should be considered with the framework of sklearn

TABLE 2: Table with grid of hyper-parameters of each model

In Figure 4 the accuracy of each model as a function of the number the features that remained after the application of mRMR is presented. It is possible to see that the performance of the models is very disappointing, as the accuracy of most of them is very close to 33%, which is equivalent to classifying examples at random.

It is clear that the number of features being suficient is not the problem, as the accuracy of the models reaches a plateau around 400 features, and increasing the number of features worsens the accuracy of the models, most likely causing them to overfit.

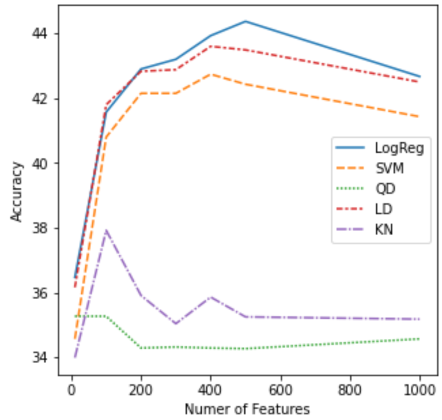

FIGURE 4: Accuracy per model as a function of number of features

To understand if the problem was related to the quality of the features, that is, to verify if the features were not being sufficiently informative about the examples, we applied a different dimensionality reduction technique to the whole set of 5000 features, Linear Discriminant Analysis. After applying this method, only two features remained, which means they can be visualised. In Figure 5 we present the examples of the dataset plotted against the two features resultant of Linear Discriminant Analysis, where the different classes are clearly separated, meaning the features returned by the pre-trained models are sufficiently informative.

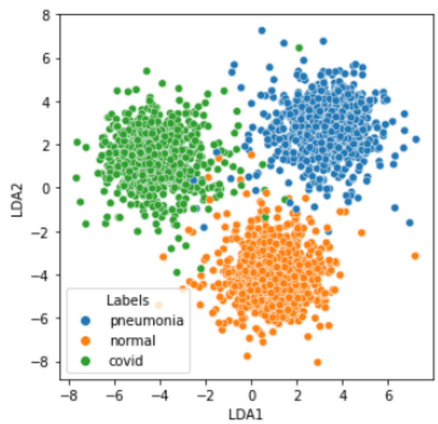

FIGURE 5: Examples as a function of features after LDA

We then tried to use Linear Discriminant Analysis instead of mRMR for feature reduction. However, similar results to the ones presented before were found. It is important to notice that when the Linear Discriminant was used to plot the previous graph, it applied to both training and validation set. However, when used with the classifiers, the Linear Discriminant Operator (within Sklearn) is fitted to the training data and only then applied to the validation dataset.

In conclusion, we haven't been able to successfully implement this approach of transfer learning with feature extraction and haven't been able to identify the reason for our insucess, as it appears the features returned by the pre-trained models are sufficiently informative.

B. TRANSFER LEARNING WITH FINE TUNING

For transfer learning with fine-tuning, we decided to follow an approach similar to that proposed in [6], where the accuracy obtained for binary classification was also very high 96.4%. In this work, the last layer of different pre-trained networks was replaced by a simpler layer that was then trained. In the end, an ensemble classifier was made with all the new networks.

The pre-trained networks that we used are the sames as the ones used in the previous section: MobileNetV2, VGG16, VGG19, DenseNet121, ResNet50V2. The last layer of each one of these networks was substituted by a layer composed of three neurons (corresponds to the number of classes, which is also three) with the SoftMax activation function. Each model has about 6000 parameters to tune. To help prevent overfitting, data augmentation was used, more specifically, random rotations and flips along the horizontal axis were applied to the images.

Before passing the images through the models, the images were scaled to the size 224x224, and were pre-processed according to the requirements of each model.

Each model was trained for 10 epochs with an adaptative learning rate and with the Adam optimizer. The training and validation accuracy in each epoch per model are presented in Figures 6 and 7.

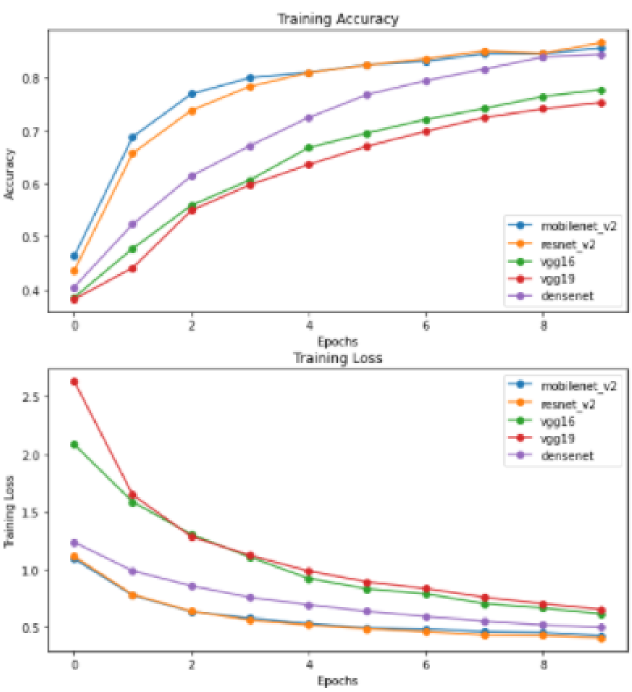

FIGURE 6: Training Accuracy and Loss

Due to time constraints, the number of epochs was limited. However, it is possible to see that for DenseNet, there was

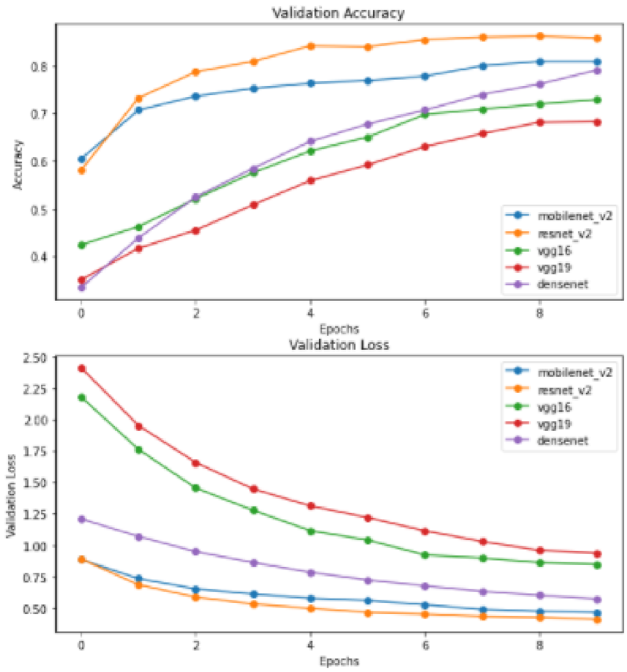

FIGURE 7: Validation Accuracy and Loss

probably still room of improvement for accuracy, as the validation accuracy was still growing. For the other models, it seems that the validation accuracy has reached a plateau and there would not be much improvement if the number of epochs was increased.

An ensemble classifier based on hard voting was built using the previsions made by these models. In Table 3 we present the metrics of each model, obtained in the validation set.

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV2	83.0	83.2	83	82.9
ResNet50V2	**86.9**	**87.3**	**86.9**	**87.0**
VGG16	73.7	75.8	73.7	73.3
VGG19	68.0	69.8	68.0	68.3
DenseNet121	80.0	81.3	80.6	80.5
Hard-Voting	86.7	87.1	86.7	86.8

TABLE 3: Table with metrics of each model in the validation set

The model with the best performance was ResNet50, surpassing the Hard Voting classifier and all the other models. This is not surprising, as ResNet was also the model which achieved the best performance in the paper that inspired this approach. The confusion matrix of this model for the validation set is presented in the following figure.

The confusion matrix of RestNet50 for the testing set is presented in Figure 9. In Table 4 the metrics of performance

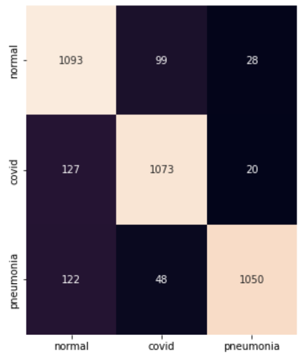

FIGURE 8: Confusion Matrix of ResNet50 for Validation Set

for the testing set are presented. The final accuracy of the model is 87.1%.

Observing the confusion matrix, it is possible to see that most errors are concerned with the Normal and Covid Classes.

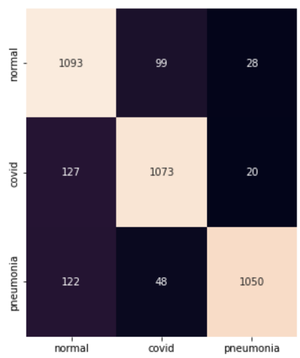

FIGURE 9: Confusion Matrix of ResNet50 for Testing Set

Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
87.1	88.3	87.9	88.0

TABLE 4: Metrics of ResNet50 for Testing Set

Comparing this model with the ones presented in literature, its accuracy, although inferior to the accuracies of the models we analysed, is close to them. Comparing solely this model with the models that also used Transfer Learning with Fine Tuning, those models achieved accuracies of 90% and 93%, which are not very far off from the accuracy we obtained.

Apart from the model, one important conclusion we reached in this work is that MobileNetV2 and RestNet50V2 seem to have the best performances for this multiclass problem.

1) Future Work

To mitigate the difference between the accuracy obtained in this work and the accuracies obtained in literature, we propose the following approaches, that could be used as future work:

- Increase the number of epochs used, especially for the DenseNet model;
- Test the models for different sizes of input images;
- Vary the number of final layers of the pre-trained models;.

One other approache that could also be interesting as future work, but which is not limited to the previous model is:

- Convert this multi-class classification problem into two binary classification problems, by starting by training a model that distinguishes images belonging to one of two classes of Healthy vs Covid+Pneumonia. Train a second model that distinguishes between Covid vs Pneumonia. If the level of similarity of the examples belonging to the classes of Covid and Pneumonia is greater than the level of similarity between examples belonging to the classes of Covid and Normal, or Pneumonia and Normal, the accuracy of the overall model could possibly be improved by creating a model that specialises in distinguishing between patients with Covid and Pneumonia.

V. CONCLUSION

The goal of this work was to develop a deep learning model to classify X-Ray images of healthy patients, patients with Covid-19 and patients with pneumonia. This task can be hard due to the fact that if the disease is caught in an early stage, it's unlikely that anything substantial will appear on the X-Ray.

Two models were trained to solve this problem. In the first model, five pre-trained CNNs were used to extract features from the images in our dataset. After feature reduction with mRMR was applied, several different linear classifiers were trained with these features. The results were disapointing. We couldn't determine the source of the problem, as it was shown, with Linear Discriminant Analysis, that the features contained relevant information about the data. In the second model the last layer of five pre-trained CNNs was substituted by a simple layer containing three neurons. An ensemble classifier based on hard voting was obtained with the CNNs. The results achieved were satisfactory. The CNN RestNet and the CNN DenseNet had the best performances. The accuracy of the modified ResNet, which was the model with the best performance, was 87.1%. This value is inferior to most values present in literature, but is is still very satisfactory. Some suggestions of future work that could be done to improve the performance of the model are: increate the humber of epochs, as due to time constraints, only 10 epochs were ran; vary the number of final layers of the pre-trained models.

VI. WORK DONE BY EACH STUDENT

We think that the work developed on this project was well divides by both of the students of the group, but, in detail, the work done by each one of us was:

Vitória Cruz: development of the model based on transfer learning with feature extraction, writing of the report

João Alegria: search of the dataset, development of the model based on transfer learning with fine tuning, formatting of the report,

REFERENCES

[1] Fan Wu., Zhao Su., Bin Yu., Chen Yan-Mei, Wang Wen, Song Zhi-Gang, Yi Hu., Tao Zhao-Wu, Tian Jun-Hua, Pei Yuan-Yuan. A new coronavirus associated with human respiratory disease in china. Nature. 2020;579(7798):265–269.

[2] Bernheim A., Mei X. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. Radiology. 2020 doi: 10.1148/radiol.2020200463.

[3] Long C., Xu H. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? Eur. J. Radiol. 2020;126:108961.

[4] M. Toğaçar, B. Ergen, Z. Cömert, F. Özyurt, A Deep Feature Learning Model for Pneumonia Detection Applying a Combination of mRMR Feature Selection and Machine Learning Models, IRBM, Volume 41, Issue 4, 2020, Pages 212-222, ISSN 1959-0318

[5] Z. Zhao, R. Anand and M. Wang, "Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform," 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2019, pp. 442-452, doi: 10.1109/DSAA.2019.00059.

[6] Chouhan, V.; Singh, S.K.; Khamparia, A.; Gupta, D.; Tiwari, P.; Moreira, C.; Damaševičius, R.; de Albuquerque, V.H.C. A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images. Appl. Sci. 2020, 10, 559. https://doi.org/10.3390/app10020559

[7] Hashmi MF, Katiyar S, Keskar AG, Bokde ND, Geem ZW. Efficient Pneumonia Detection in Chest Xray Images Using Deep Transfer Learning. Diagnostics (Basel). 2020;10(6):417. Published 2020 Jun 19

[8] Okeke Stephen, Mangal Sain, Uchenna Joseph Maduh, Do-Un Jeong, "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare", Journal of Healthcare Engineering, vol. 2019, Article ID 4180949, 7 pages, 2019. https://doi.org/10.1155/2019/4180949

[9] Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed. 2020;196:105581. doi:10.1016/j.cmpb.2020.105581

[10] Islam MZ, Islam MM, Asraf A. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. Inform Med Unlocked. 2020;20:100412

[11] Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med. 2020;43(2):635-640. doi:10.1007/s13246-020-00865-4

[12] Wang, Linda, Zhong Qiu Lin, and Alexander Wong. "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images." Scientific Reports 10.1 (2020): 1-12.

[13] Hemdan, Ezz El-Din, Marwa A. Shouman, and Mohamed Esmail Karar. "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images." arXiv preprint arXiv:2003.11055 (2020).

[14] Asraf, Amanullah; Islam, Zabirul (2021), "COVID19, Pneumonia and Normal Chest X-ray PA Dataset", Mendeley Data, V1, doi: 10.17632/jctsfj2sfn.1

[15] Samritika Thakur, Aman Kumar, X-ray and CT-scan-based automated detection and classification of covid-19 using convolutional neural networks (CNN), Biomedical Signal Processing and Control, Volume 69, 2021, 102920, ISSN 1746-8094,

[16] Shazia, A., Xuan, T.Z., Chuah, J.H. et al. A comparative study of multiple neural network for detection of COVID-19 on chest X-ray. EURASIP J. Adv. Signal Process. 2021, 50 (2021)