Screening for COVID-19 using Chest X-Ray images and ML techniques

Dorra El Mekki
elmekk01@ads.uni-passau.de
University of Passau

Mohamed Abdallah Zaghlol Abdo
abdo01@ads.uni-passau.de
University of Passau

Muhammad Hadeeq
hadeeq01@ads.uni-passau.de
University of Passau

Alok Chauhan
chauha02@ads.uni-passau.de
University of Passau

ABSTRACT

The novel coronavirus SARS-CoV-2 led to the outbreak of a respiratory disease known as COVID-19. After the initial wave of COVID-19, there is an urgent need for faster and cheaper tests. COVID-19 virus affects the respiratory system and Chest X-Ray (CXR) is one of the key imaging methods for identifying the corona virus since it is cheaper and faster to obtain, compared to PCR tests. Nevertheless, it remains difficult for doctors to diagnose COVID at an early stage based on CXR images only. For this reason, This project aims to enhance prognostic predictions with the Chest X-Ray dataset. In this project, pre-trained deep learning models (such as VGG, ResNet, InceptionV3, ResNext) are used to classify COVID-19 positive cases among other diseases. Deep neural network requires a large dataset. Thus, it overfits on the CXR dataset. For this reason, we propose a two-phase approach: A pre-trained deep learning model extracts the high-level features. A first classifier divides CXR images into two main groups: healthy patient or patient with pneumonia. If the result falls in the "normal" category, the process ends. However, if the result is classified as pneumonia, then a second binary classification is carried out to split pneumonia cases into COVID-19 or non-COVID. Lastly, we compare the results of pre-trained models to the proposed two-phase approach and, on a public dataset, the two-phase approach, built with Resnet and SVM outperformed among others.

1 INTRODUCTION

As of May 03, 2020, there were total of 3,524,429 confirmed cases with 2,144,038 active cases and 247,838 deaths in more than 187 countries across the globe due to COVID-19[1]. It was first detected in Wuhan, China by the end of December 2019, and soon afterwards it has spread to almost every country. The outbreak was declared a public health emergency of international concern by the World Health Organization on 30 January, and a pandemic on 11 March. Common symptoms include fever, cough, tiredness, shortness of breath, loss of odor and taste. The main mode of transmission is through breathable droplets expelled from an infected person's mouth or nose and inhaled by a healthy person during close contact, sneezing, or talking. The exponential growth of COVID-19 cases is caused by their high transmissibility. In early studies, an average number of secondary cases generated from a single infectious case was estimated at 2.2.

COVID-19 testing is crucial for slowing down infection rates and relieving congestion in hospitals for emergency departments and intensive care units. The basic COVID-19 tests are called PCR (Polymerase chain reaction) tests which look for the presence of a certain infection's antibodies. But this analysis has a few limitations. Pathogenic laboratory testing takes time, with significant false-negative findings [15]. In addition, large-scale implementation of costly COVID-19 testing cannot be affordable by many countries. Therefore, using artificial intelligence & machine learning techniques is helpful. In this context, we are suggesting an architecture combined of two binary classifiers in order to correctly classify Covid-19 affected people. This solution will help doctors for triage process.

By dividing this project in four phases as shown in table 1, each member will focus on one of the four phases starting from defining the problem description & pipeline and then preprocessing the data with some of the advance methods such as data augmentation to further actual implementation and evaluation.

Table 1: DSL 2020 team members

Phase	Name
I)	Muhammad Hadeeq
II)	Dorra El Mekki
III)	Mohamed Abdallah Zaghlol Abdo
IV)	Alok Chauhan

2 PROBLEM STATEMENT

Our main focus is to facilitate the testing phase of the virus by using Chest X-Ray images. In this context, a considerable amount of research work has been devoted. However, our suggested approach is a combination of two binary classifiers. To conclude, our research question is the following: Is it possible to achieve better accuracy than existing state-of-the-art models by using hierarchical classification technique?

During this research work, a public dataset of CXR images were used. As a first approach, pre-trained deep learning models (VGG, ResNet, InceptionV3, ResNext) used to identify positive cases of COVID-19 among other diseases. As a second approach, we applied our own methodology, a deep learning model implemented to first

[1] https://coronavirus.jhu.edu/map.html

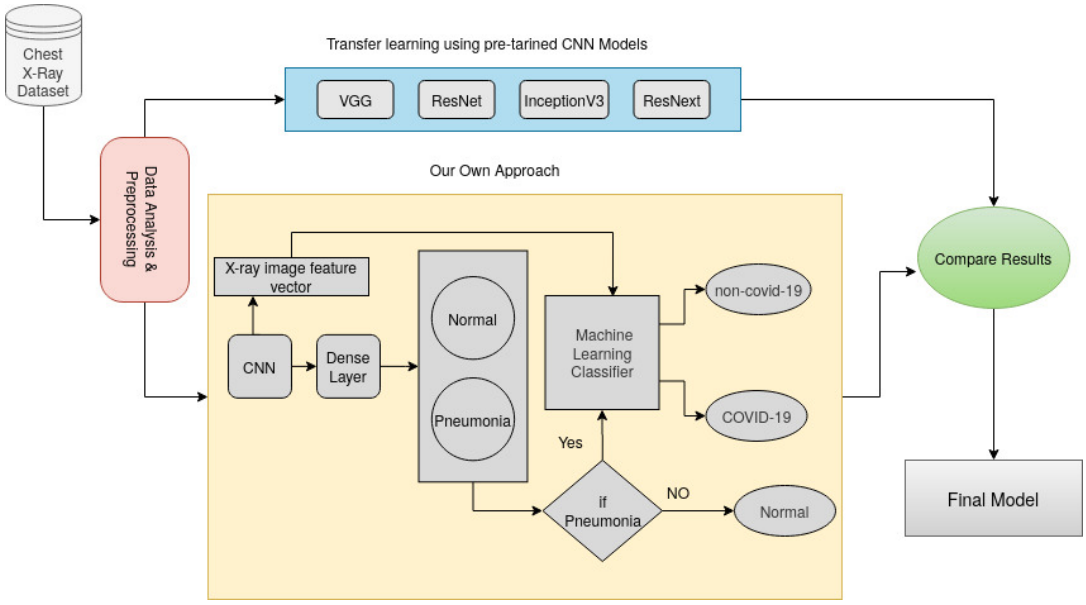

Figure 1: Pipeline

identify normal and pneumonia from CXR images, which further identified if the pneumonia is caused by COVID 19 or other disease. The focus of this project is to provide COVID-19 rapid detection method by Chest X-Ray images with pre-trained and deep learning model methodologies. Also, to select the approach that is giving the better results and has the highest accuracy and the lowest false negatives as particular concern.

3 RELATED WORK

In a hierarchical classification, COVID-19 is considered as a virus. Moreover, it is causing pneumonia. Hence, the possibility to detect COVID-19 using Chest X-Ray (CXR) images. In medicine, It is a known fact that CXR does not perform a thorough diagnosis. Indeed, doctors that rely on CXR can detect only 69% of COVID-19 cases. In contrast, RT-PCR [2] has an accuracy of 91 %. However, CXR tested positive in the 9% remaining cases. This is due to the early stage of this infectious disease, where the coronavirus exists only in the peripheral area of the lungs. Which makes the RT-PCR test not capable of detecting it. As a solution, a pulmonary examination may be possible, however, the latter is hurtful for the patients. Hence, the recourse to the X-Ray images and the use of both tests since they are complementary. In this context, the main target of this study is to use chest X-Ray images for patient triage in order to manage the priority of patients" treatments.

After discussing the approach of classifying all images to normal, bacterial, non-COVID viral infection and COVID-19 viral infection with experts in the medical field, we came out with the fact that viral and bacterial infections have very similar lesions in the lungs. However, the Coronavirus involves specific lesions that can be distinguished from other infections. Hence, instead of classifying

[2] Real-time reverse transcriptase-polymerase chain reaction (RT-PCR) is the main technique for COVID-19 diagnosis.

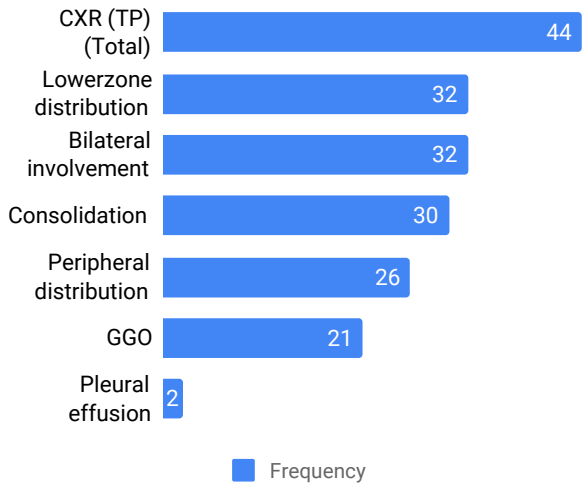

Figure 2: Most frequently observed distribution patterns of COVID-19 [16].

X-Ray images to normal, bacterial, non-COVID viral infection and COVID-19 viral infection. It is easier to classify it in the first step to normal and pneumonia. Then, in a second step subclassify pneumonia class into COVID-19 viral infection and others. According to experts, differentiating between bacteria and viruses can be based on lesions localization. If the lesions are centralized then it is more probably a bacterial infection, however, if the lesions are spreading out in the bilateral regions and the base region of the lungs then it is more likely to be a COVID-19 infection [10]. The radiological findings are confirmed again in this study [16]: COVID is detected

in the bilateral consolidation of the lungs in 47% of the cases. 33% of the patients are affected in the base of the lungs, named ground-glass opacification (GGO). Hence the importance of defining the lung contour using semantic segmentation [6]

As a matter of fact, relying only on a CNN algorithm may lead to the wrong classification since it learns motifs independently from its location, and in our case, the lesion"s location is one of the key features.

COVID-19 is a recent infectious disease. Available datasets [11] [14] [3] have a limited size. One way to avoid overfitting is to rely on attention models [6] : the image is divided into small patches that are analyzed separately. The final answer is decided by majority voting.

4 PIPELINE

Considering the complexity of the project we divided this project into the following four phases:

(1) Data analysis & preprocessing
(2) Transfer learning using pre-trained CNN models
(3) Our own approach
(4) Compare results & finalizing model

All these four phases are explained in the following sections with abstract details.

4.1 Data analysis & preprocessing

Each image has different parameters from various machines and centers, i.e. hue , saturation, contrast and scale. Therefore preprocessing is important to standardize the images in a uniform format. Images will be cropped at the first preprocessing stage to clear boundaries to ensure that the photos only contain lungs.

4.1.1 Data Cleaning. The used dataset[3] is collected from two different sources. The first dataset[4] has the dominant number of normal X-Ray images and also images with pneumonia disease. This data will be useful for the first classification phase (Normal vs Pneumonia). However, it contains only 56 Covid-19 cases. The second dataset [2], contains numerous subclasses of pneumonia diseases. From both datasets, we have collected a total of 373 Covid-19 cases, five times more than the collected data with only the first dataset.

Merging both datasets raised a new issue: duplicated images were found. In order to avoid overfitting, eliminating these images is a must. First, duplicate images with the same md5 hash are eliminated. Secondly, comparing the similarity between images allows us to identify images that are cropped or resized. Note that this analysis showed that some images with different names but the same content exist in both train and test set even from the same source of data.

4.1.2 Data Selection. After cleaning the data and removing duplicate images using the CCleaner tool, we selected the target data by eliminating unlabeled images. Figure 3 shows the data cleaning process in numbers. Respectively, 25% and 30% from dataset 1 and dataset 2 were removed. To conclude, 4407 images out of 5910 were

[3] Our dataset is publicly available on: https://www.kaggle.com/dorraelmekki/chestxray-images-for-covid19-the-largest-dataset
[4] Source of the 1st dataset: https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset

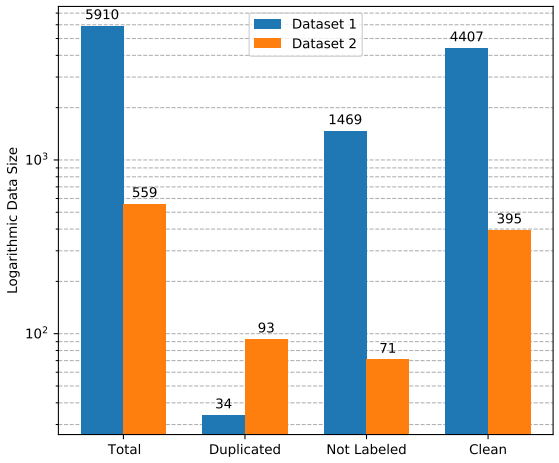

Figure 3: Data statistics

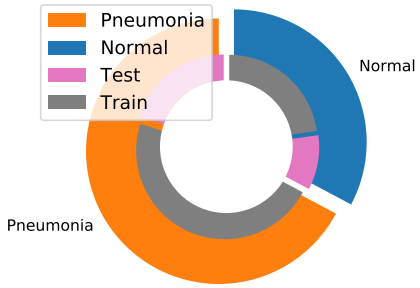

Figure 4: Data distribution of 1st phase of classification

selected from the 1st dataset and 395 out of 559 images from the 2nd dataset to finally get 4802 images.

4.1.3 Data Preparation & Description. A standard way to organize the image data is to first sort them by dataset (e.g. Train, test, and validation). Then, create subdirectories; one for each class under the dataset directory. In our case, we created a root directory that contains train and test folders. Due to the limited data, cross-validation will be applied as a substitute for the validation set. The Train-Test Split ratio is 7:3.

In order to have the same percentage of data points of each class in train and test sets, we applied the stratified sampling.

According to our pipeline (Figure 1), we have two classification phases. Normal vs Pneumonia, then in a second step; COVID-19 vs Non-COVID-19. Figure 4 and 5 represent the data distribution of the 1st and the 2nd phase of classification. The classes are unbalanced. Thus, data augmentation and upsampling techniques will take place in order to deal with biased data and avoid overfitting. The characteristics of the data are summarized in table 4 and table 3. To conclude, the labels were divided into two main classes: normal and Pneumonia and two subclasses; Covid-19 and non-COVID-19.

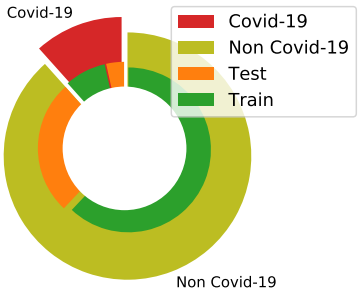

Figure 5: Data distribution of 2nd phase of classification

Table 2: 1st phase classification

Classes	Train	Test	Total
Normal	1094	480	1574
Pneumonia	2264	964	3228
Total	3358	1444	4802

Table 3: 2nd phase classification

Classes	Train	Test	Total
Covid-19	262	113	375
Non-COVID-19	2002	851	2853
Total	2264	964	3228

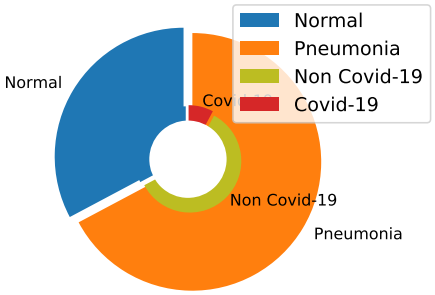

Figure 6: Data distribution

In the case of having Pneumonia disease, we move to the second step to identify its origin. If it is a pneumonia disease caused by COVID-19 infection or not. For a global visualization of the data distribution, see Figure 6.

Distinguishing different classes of pneumonia diseases (e.g Bacterial pneumonia, SARS-CoV, ARDS) is very challenging even for experienced radiologists [6]. Thus, we assigned the same class "non Covid-19" for bacterial pneumonia, SARS-CoV, 'ARDS'. All these diseases have similar radiologic features.

4.1.4 Data Processing. Figure 7 is the example of positive CXR for COVID-19 with radiographic findings. The green and yellow

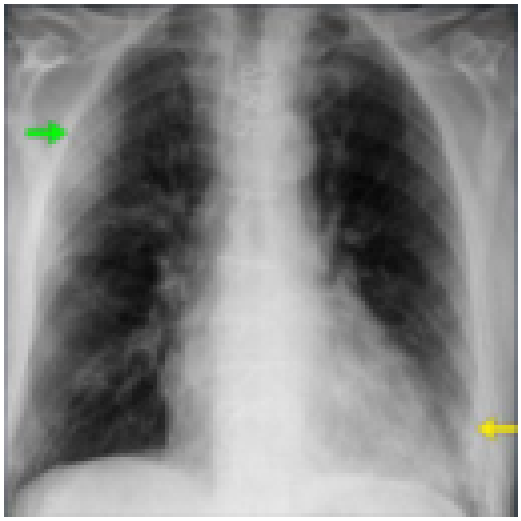

Figure 7: COVID-19 positive CXR

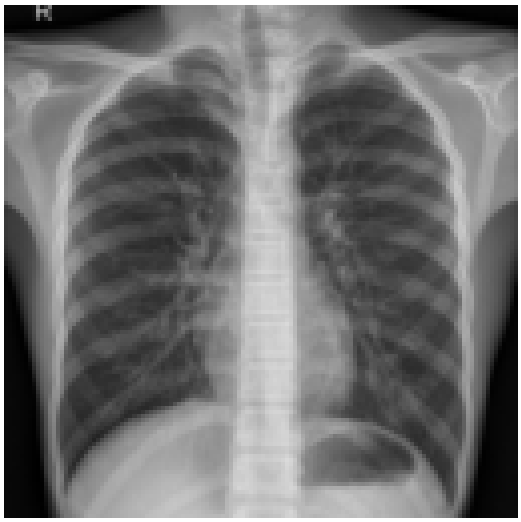

Figure 8: Normal CXR

arrows are pointing to the white space in both lungs, particularly the lower lobes, are depicting COVID-19 effects. Furthermore, if the lesions are centralized, it is more likely to be a bacterial infection, but if the lesions spread in bilateral regions and the base area of the lungs, COVID-19 infection is more likely to occur [10]. COVID is detected in the bilateral consolidation of the lungs in 47% of the cases [6].

CXR images are collected from different resources. This means images have different characteristics such as compression type or image size. Thus, resizing and normalizing the data is an essential step.

4.1.5 Data augmentation. Training deep artificial neural networks requires a large dataset. However, it is not guaranteed to collect huge data as it may be expensive and sometimes very rare. As a

solution for this issue, it is possible to generate artificial data for training, thanks to data augmentation methods [12].

From existing data augmentation techniques shifting and cropping are used. As mentioned in the reference paper [6], applying these techniques may fluctuate the performance by approximately ±1.01%. Rotated images was not applied since we cannot find rotated CXR images in real data.

4.2 Transfer learning using pre-trained CNN Models

During the implementation phase, we used different models in order to compare the result and consider the best one. We initialized the network with pre-trained weights, where knowledge is transferred from a different domain and task. According to our pipeline in Figure 1, we first fine tune the model on training data then evaluate it on test data and finally pass the results to comparing phase. The models which we have investigated (VGG, ResNet & Inception V3) are the state of art models in image classification and are well trained on an image dataset.

In our use case, first, we transformed the images by cropping, horizontally flipping and normalizing then putting them into batches and feed forwarding every batch of images into the network. We also made a change in the last layer of the network and froze initial CNN layers. Instead of keeping 1000 output classes in the last layer, we kept three classes which are 'Normal', 'non-COVID-19' and 'COVID-19' to get the predictions. Moreover, we followed the same procedure to train ResNet, Resnext, Inception_v3 to check the performances of these architectures on the same dataset. Among all the models we have got the best accuracy using the ResNet pretrained model. In VGG-16, Resnet, Resnext, Inception_v3 we have used ReLU as an activation function, stochastic gradient descent as an optimizer, and cross-entropy as a loss function.

4.2.1 ReLU. Rectified Linear Unit activation is a non-linear activation function[5] and most widely used in deep learning applications such as computer vision, speech recognition, and deep neural network, etc. The mathematical expression of this function is in Formula 1. That means if the x value is negative, then it will return zero. If the value of x is positive, it will always return the input value. ReLU has a lower run time compared to other activation functions. It does not return backpropagation errors or in other words, vanishing gradient problems will not occur and weights will converge easily.

$$\mathrm{ReLU}(x) = \max(0, x) \tag{1}$$

4.2.2 Stochastic gradient descent. SGD [9] is an optimization algorithm used to reduce the loss (mismatch between y_pred and y). While trying to find the derivative of the loss considering only one data point at a time which basically means at every epoch it is taking only one data point and calculating y_pred and while backpropagating updating the weights. There are variants of SGD such as Mini batch SGD, Batch SGD. In the case of Mini batch SGD, it will be considering only k data points, where k is always less than the total number of data points. However, in the case of gradient descent, it will be using all the millions of samples for completing one epoch and it has to be repeated for every epoch till it reaches the minima.

4.2.3 Cross-entropy loss. We have used the cross-entropy loss function 2 to compute the difference between predicted labels and actual labels[6].

$$CE = -\sum_{i}^{C} t_i log(f(s)_i) \tag{2}$$

Where t_i is the true labels and $f(s)_i$ is the predicted label.

4.3 Our approach

In this approach, the method has been divided into two independent classifiers. As per the pipeline (Figure 1), first we pass the CXR image to the pretrained CNN model to obtain the feature vector and the first phase classification result (normal or pneumonia). If the result is classified as pneumonia, then we pass the feature vector from the image to machine learning algorithm (e.g. support vector machine or logistic regression), which will be independently trained to classify between (Covid-19 and non-covid-19). Pseudo code of this approach is mentioned in Algorithm 1.

4.3.1 First phase of classification (normal, pneumonia). After observing the results of pretrained models (Table 4), we chose to mimic the same CNN architecture of Resnet-18 [1] and train it on our dataset, the difference here is that we will re-train all the layers without freezing any of them to detect if the image is normal or pneumonia. During this project, we trained the model using the following hyperparameters; epoch = 25, batch_size = 8 and stochastic gradient descent as an optimizer with 0.01 learning rate and cross entropy loss as a loss function. The results are shown in Appendix A.

4.3.2 Second phase of classification (Covid19, non Covid19). When we trained the the CNN model, we preserved the feature vector of Covid and non-Covid images to perform the training of the second phase classification. In this phase, we trained more than one classifier, we compared the results and chose the best one in terms of accuracy and F1-score. We investigated SVM [13], Logistic Regression [4], decision trees [7], and FNN. We observed that SVM outperformed all the previous models by 10% of accuracy. We performed a grid search technique to find the best set of hyperparameters for C and gamma 'C': 100, 'gamma': 0.001.

5 COMPARING RESULTS & FINAL MODEL

Table 4: classification reports for all investigated models

Models	Precision	Recall	F1-Score	Accuracy
VGG-16	79%	83%	80%	80%
Resnet	84%	85%	84%	84%
Resnext	76%	78%	77%	77%
Inception_v3	74%	69%	70%	74%
Resnet + SVM (our approach)	**89%**	**90%**	**89%**	**94%**
VGG + SVM	84%	85%	85%	87%

[5] https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

[6] https://gombru.github.io/2018/05/23/cross_entropy_loss/

Algorithm 1 Pseudocode of our approach

```
for first_image to last_image do
    result ← first_phase_classifier(image)
    feature_vector ← get_feature_vector(image)
end for
if result = pneumonia then
    second_classifier_results ← sec_clf(feature_vector)
    if second_classifier_results = Covid_19 then
        return Covid
    else
        return non_Covid
    end if
else
    return Normal
end if
```

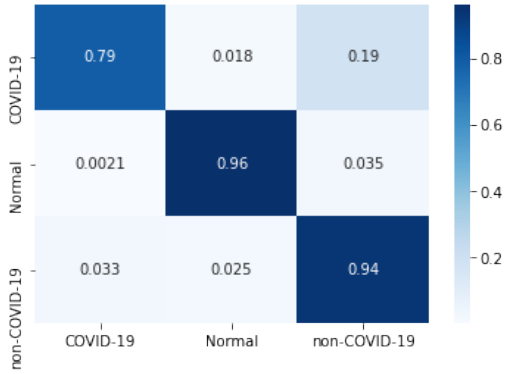

Figure 9: Confusion matrix of Resnet+SVM

The experiments have been performed with different model architectures (VGG-16, Resnet, Resnext, Inception_v3) to check the accurate results.

Figure 9 represents the confusion matrix obtained by the combination of Resnet and SVM. Moreover, the analysis in table 4 show the Precision, Recall, F1 score and accuracy of the experiment conducted in both phases. As observed, our approach (Resnet+SVM) has outperformed all other models in terms of accuracy, precision, recall and F1 score.

6 DISCUSSION

We propose a hierarchical classification approach to improve classification performance. Deep convolution neural networks require a large amount of data to process in order to perform well. In our case, we only had 4802 images. Therefore, we have used ML classifiers since they perform well on a small dataset.

We compared our results with state of the art models and we found that our approach (Resnet+SVM) outperformed the model introduced in [5] by 4%. Furthermore, to check the consistency in the performance, we trained all the models multiple times and found that results are quite stable. This study has potential limitations. Hierarchical classification is engendering the cumulative error. In other words, if the CNN model misclassified any of the given classes then the second phase of classification will be affected. Thus, the final result will be wrong.

7 CONCLUSION

The main idea of this project is to detect Covid-19 from chest x-ray images by using ML techniques and check whether is it possible to achieve a better performance than existing state-of-the-art models. In order to perform the experiment, the data has been collected from two different sources, concatenate into one and select the target data by eliminating unlabelled images. We used 3 cross-validations, as a workaround for a small no representative validation set. Then we applied a train-test split ratio of 7:3. Experimental results of the first approach (one phase classification) showed that the ResNet model performed the best in classifying “Covid-19”, “Normal”, and “non_Covid-19” class. Therefore, we employed ResNet (for feature extraction) in our second approach. Our two-phase approach classifies the CXR images into “Normal” and “Pneumonia”. In the second phase, the pneumonia class is sub-classified into “Covid-19” and “non-Covid-19” classes. This hierarchical classification enabled us to work with two datasets that are more balanced. Hence, we obtained better result, compared to the first approach. Lastly, stacking ResNet and SVM outperformed all other combinations of ML models.
For future work, we aim to explore the domain-specific model like chestXnet [8]. As it is a 121 layer deep convolution neural network, given the images of chest x-ray, it outputs the probability of the class along with its heat map. Furthermore, we aim to use up-sampling technique to mitigate the bias in the data.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to our supervisor, Mr. Julian Stier as well as Prof. Dr. Michael Granitzer for their consistent support and constructive remarks.
Our sincere thanks and gratitude go to Dr. Ines Belaid, Dr. Fethi Belaid, and Dr. Naouel Belaid for sharing with us their knowledge and expertise.
We extend our thanks to Mr. Mohamed Karim Belaid for his comments and thoughtful efforts to improve the quality of the work.

REFERENCES

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
[2] P. Morrison J. P. Cohen and L. Dao. 2020. COVID-19 image data collection. *arXiv 2003.11597* (2020). https://github.com/ieee8023/covid-chestxray-dataset
[3] Stefan Jaeger, Sema Candemir, Sameer Antani, Yì-Xiáng J Wáng, Pu-Xuan Lu, and George Thoma. 2014. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery* 4, 6 (2014), 475.
[4] David G Kleinbaum, K Dietz, M Gail, Mitchel Klein, and Mitchell Klein. 2002. *Logistic regression*. Springer.
[5] Arpan Mangal, Surya Kalia, Harish Rajgopal, Krithika Rangarajan, Vinay Namboodiri, Subhashis Banerjee, and Chetan Arora. 2020. CovidAID: COVID-19 Detection Using Chest X-Ray. *arXiv preprint arXiv:2004.09803* (2020).
[6] Yujin Oh, Sangjoon Park, and Jong Chul Ye. 2020. Deep Learning COVID-19 Features on CXR using Limited Training Data Sets. *arXiv preprint arXiv:2004.05758* (2020).
[7] J. Ross Quinlan. 1987. Simplifying decision trees. *International journal of man-machine studies* 27, 3 (1987), 221–234.
[8] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. 2017. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225* (2017).

[9] Sebastian Ruder. 2016. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747* (2016).
[10] Sana Salehi, Aidin Abedi, Sudheer Balakrishnan, and Ali Gholamrezanezhad. 2020. Coronavirus disease 2019 (COVID-19): a systematic review of imaging findings in 919 patients. *American Journal of Roentgenology* (2020), 1–7.
[11] Junji Shiraishi, Shigehiko Katsuragawa, Junpei Ikezoe, Tsuneo Matsumoto, Takeshi Kobayashi, Ken-ichi Komatsu, Mitate Matsui, Hiroshi Fujita, Yoshie Kodera, and Kunio Doi. 2000. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology* 174, 1 (2000), 71–74.
[12] Luke Taylor and Geoff Nitschke. 2017. Improving deep learning using generic data augmentation. *arXiv preprint arXiv:1708.06020* (2017).
[13] Simon Tong and Edward Chang. 2001. Support vector machine active learning for image retrieval. In *Proceedings of the ninth ACM international conference on Multimedia*. 107–118.
[14] Bram Van Ginneken, Mikkel B Stegmann, and Marco Loog. 2006. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Medical image analysis* 10, 1 (2006), 19–40.
[15] Shuai Wang, Bo Kang, Jinlu Ma, Xianjun Zeng, Mingming Xiao, Jia Guo, Mengjiao Cai, Jingyi Yang, Yaodong Li, Xiangfei Meng, et al. 2020. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *MedRxiv* (2020).
[16] Ho Yuen Frank Wong, Hiu Yin Sonia Lam, Ambrose Ho-Tung Fong, Siu Ting Leung, Thomas Wing-Yan Chin, Christine Shing Yen Lo, Macy Mei-Sze Lui, Jonan Chun Yin Lee, Keith Wan-Hang Chiu, Tom Chung, et al. 2020. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiology* (2020), 201160.

A APPENDIX

In this section, we report the loss and accuracy of the pre-trained models. From top to bottom: Vgg-16, Ressnet, Resnext, our approch (ResNet+SVM).

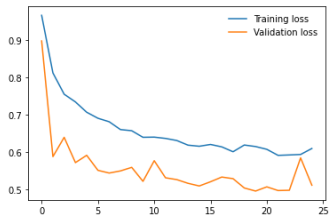

fig1: Vgg-16 loss

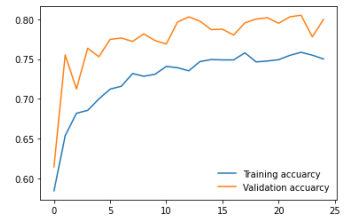

fig2: Vgg-16 accuracy

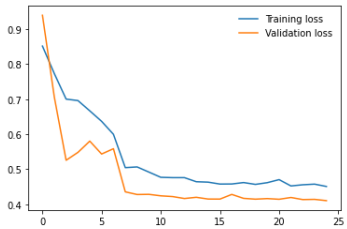

fig3: Resnet loss

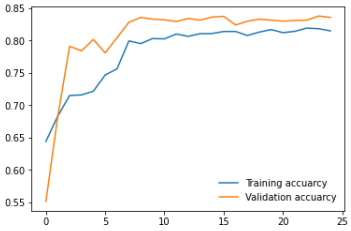

fig4: Resnet accuracy

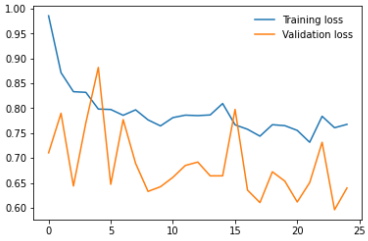

fig5: Resnext loss

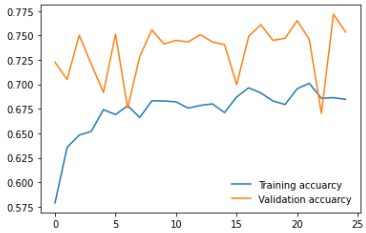

fig6: Resnext accuracy

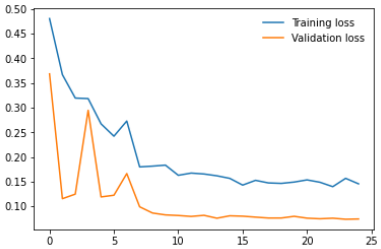

fig7: ResNet+SVM loss

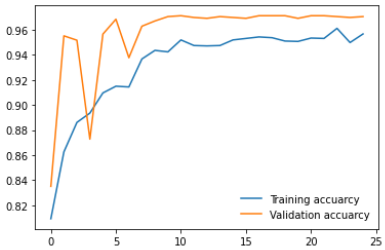

fig8: ResNet+SVM accuracy

Figure 10: Loss and accuracy curves of the pre-trained models and our hierarchical approach.