Covid19 Diagnosis: Comparative Approach Between Chest X-Ray and Blood Test Data

Ali Emre Öztaş
Department of Electronics and Communications Engineering
Istanbul Technical University
Istanbul, Turkey
oztas18@itu.edu.tr

Dorukhan Boncukçu
Department of Electronics and Communications Engineering
Istanbul Technical University
Istanbul, Turkey
boncukcu17@itu.edu.tr

Ege Özteke
Department of Electronics and Communications Engineering
Istanbul Technical University
Istanbul, Turkey
ozteke18@itu.edu.tr

Mahir Demir
Department of Electronics and Communications Engineering
Istanbul Technical University
Istanbul, Turkey
demirma18@itu.edu.tr

Arzu Mirici
Medicine Faculty
Çanakkale Onsekiz Mart University
Çanakkale, Turkey
arzumirici@comu.edu.tr

Pınar Mutlu
Medicine Faculty
Çanakkale Onsekiz Mart University
Çanakkale, Turkey
pinarmutlu@comu.edu.tr

2021 6th International Conference on Computer Science and Engineering (UBMK) | 978-1-6654-2908-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/UBMK52708.2021.9558969

***Abstract*—The Covid-19 virus has made a major impact on the world and is still spreading rapidly. A reliable solution to prevent further damage, early diagnosis of coronavirus patients are incredibly important. While chest X-Ray diagnosis is the easiest and fastest solution for this, an average radiologist has only a 75% to 85% accuracy when evaluating X-Ray data, thus it is desirable to achieve an accurate artificial network for this. Throughout this study, chest X-Ray data and blood routine test data are utilised and compared. X-Ray data consists of 5000 chest X-Ray images which are gathered from an open-source research and from a local hospital in which both have anonymous data. The blood test results were also taken from the same hospital. For the chest X-Ray diagnosis we utilised two of the popular convolutional neural networks, which are Resnet18 and Squeezenet and concluded that Resnet18 provided slightly more accurate results, while both having almost 98% accuracy. For blood test diagnosis, a feed-forward multi layer neural network was used. Even though it was worked on an insufficient dataset, 72% accuracy was obtained, thus making it a feasible option for further research. Hence, we concluded that in general chest X-Ray diagnosis is preferable over routine blood test diagnosis and the usage of AI yields better approximate results than humans.**

Keywords—covid19 diagnosis, supervised learning, chest x-ray, blood test, transfer learning

I. Introduction

With the outbreak of coronavirus in 2019 and its sudden spread throughout the world, it can be agreed upon that many of the countries' healthcare facilities were not prepared for such a pandemic. Currently there are more than 107 million cases and 2 million deaths caused by this outbreak. While some countries were more strict and careful with their precautions and thus had more control over the rising number of patients, health workers of under-developed or heavily populated areas are constantly being overworked. Moreover, as Covid19 cases increase, the costs for tests are taking a toll on both governments and its citizens.

Currently tests are done by X-Ray scans, Computer tomography and PCR-ICT tests. Among these testing methods, PCR-ICT is being used frequently due to its simplicity, speed and accuracy. However, the specificity of this method differs between 95% and 99.7% and sensitivity differs between 70% and 98% among different models of test kits. According to WHO, test methods that have the sensitivity above 80% and specificity above 90% are labeled as acceptable besides the ones that have the sensitivity above 90% and the specificity above 99% are labeled as desired[5].

Furthermore, considering the errancy in real world radiologists' performance relating to their experience, tomography scans are usually preferred for better accuracy, even though patients are exposed to radiation more than x-ray scans. With chest x-ray scans alone, when compared, radiologists with experience in the field for more than 10 years versus those with less, their reader performance are 83.7% and 76% consecutively[2].

As a result, many researchers are now trying to model networks, which can accurately diagnose patients without complete supervision, and automating the process as a means of reducing the workload on healthcare workers.

In order to diagnose Covid19, medical image processing techniques that involve deep learning are used. However, dataset size is generally a significant issue for medical image processing applications. Therefore transfer learning and pre-trained CNN models are often used to overcome this problem. In recent researches, these methods are utilised and provided promising results as follows.

Abbas et al. validated a deep CNN called Decompose, Transfer and Compose(DeTraC) for image classification with different pre-trained CNN models. The reason for that was to test and train non-homogenous and irregular data more accurately. They obtained the best accuracy with the VGG19 model(93.1%. accuracy, 100% sensitivity and 85.18% specificity). However, as their dataset is too small (105 covid | 80 non-covid | 11 SARS) the results may be overfitted and are not reliable[1].

Also in their research, Wang et al. made their own tailored open source deep neural network design and obtained good accuracy results[16].

Narin et al. used five pre-trained CNN models and obtained the best result with ResNet50. In the research, they trained three different models for three different datasets. Their model does not classify between covid and non covid patients. In other words this research suffers from the fact that these models are binary classification models and not suitable for medical diagnosis. Also some of their models

have 100% sensitivity or specificity which makes these results questionable for overfitting[13].

Alazab et al. utilized the VGG16 model for X-Ray diagnosis and got 99% F-score[3].

Minaee et al. used four different pre-trained CNN models for classification and got the best accuracy from ResNet18 and SqueezeNet. "Moreover , they have implemented heatmaps for indicating possible infected regions" . "Their research has a flaw of taking their covid patients dataset from only one source which hinders generalization." We used this approach for our research since it is relatively simple to realize and provides good accuracy scores[12].

Also, blood routine tests can also be used as a method for initial screening of Covid-19 for some certain benefits. PCR tests are rather inefficient in time compared to hemograms as the tests consume a long time period of 3-4 hours to provide results and are more expensive. Furthermore, qualified staff in laboratories with advanced equipment and certain materials are needed[17]. On the other hand, blood routine tests are delivered quickly with a hematology analyzer in 30 minutes to 2 hours[25].

Therefore, some researchers have also studied the feasibility of routine blood tests for Covid-19 diagnosis.

Brinati et al. developed two different models, with accuracies between 82%-86% and sensitivities between 92%-95%[17].

Banerjee et al. modelled networks comparing patients in hospital wards and those who are not admitted to a hospital. For the former model, they acquired circa 90% accuracy and sensitivity in between 82% to 92%. The latter models predicted accuracies 82%-87% and sensitivities 43%-65%[18].

The accuracy of the ensemble model of Aljame's et al. is 99.88% even though the dataset from Albert Einstein Hospital in Brazil has too many missing values. So, according to these results, the correctness of models is questionable[19].

There are four main contributions of this paper. Firstly, this paper exhibits the performance comparison of covid prediction with X-Ray and blood test data. Secondly, we developed a new model to predict from blood test data and adjusted existing models to obtain better results. Moreover, we collected our dataset from multiple sources to make our model more reliable and generalized. Lastly, we trained our SqueezeNet and ResNet18 models into two different datasets for X-ray diagnosis. Models trained with the first dataset classify between the covid patients and non-covid patients including every type of disease. Models trained with the second dataset makes binary classification between covid patients and healthy people.

This paper is organized as follows. The data contents and its preparation is explained in Section 2. We have mentioned the required theoritical information for the methods used in this paper in Section 3. Our models are explained in Section 4. The experimental results and evaluation scores are displayed and commented on in Section 5. Paper is concluded in Section 6.

II. DATA

A. Chest X-Ray Dataset

First of all, we started our work with the data set that Servin Minaee et al. elaborated for their work[12]. For Covid images, this set used Chest X-Ray-Dataset collected by Joseph Paul Cohen. Cohen's dataset is gradually updated and we used the one that Minaee et al. used for their work[8]. Then after consulting a radiologist they took only anterior-posterior images with the clear sign of the Covid for the diagnosis purposes. At the end, this set consists of 184 Covid images. For non-covid dataset we used Chex-Pert dataset[12] as Shervin Minaee et al. used. This dataset contains both X-Ray images of healthy people and patients of various diseases with respect to 13 categories. Grand total is 5000 images for non-covid dataset. We also get 151 Covid and 35 Non-Covid images without any personal information from Canakkale Onsekiz Mart University Chest Diseases Department. The table below shows the exact number of images that we used as train and test datasets.

TABLE I. DATA NUMBERS for COVID and NON-COVID PATIENT X-RAYS

	Test	Train
Covid	145	190
Non-Covid	3010	2025

TABLE II. DATA NUMBERS for COVID and NO-FINDING PATIENT x-RAYS

	Test	Train
Covid	100	84
No-Finding	1700	198

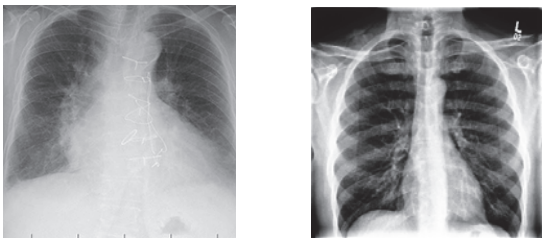


Fig. 1. Some examples for X-ray (The left image belongs to Covid-19 patient and the right one belongs to NonCovid-19 patient.)

Our data is imbalanced which means we have significantly more X-Ray image of non-covid than covid since this pandemic is recent and reaching the X-Ray data is not allowed generally.

As we have taken data from two different sources, the data formats were different. We needed to preprocess them in order to get consistent training. Therefore, we resize all images to 224x224 px sizes. At the same time some of the images were containing indicator labels and some medical terminological symbols. We removed them for a smooth process. Lastly, as we had two different datasets, we merged them by separating the set that we got from Canakkale Onsekiz Mart University into %30 test and %70 train sets.

B. Blood Test Dataset

To make a feasibility study and a comparison between the chest x-ray diagnosis method, we obtained a labelled dataset from Canakkale Onsekiz Mart University Chest

Diseases Department again. The data includes the blood routine test data from Covid-19 suspected patients. The patients' data does not contain any personal information. For Covid-19 diagnosis from blood test, we only selected necessary parameters since some parameters are not directly related to Covid-19 and can cause misleading results. The data was made usable after removing the faulty rows and the following table, illustrates the latest dataset.

TABLE III. BLOOD TEST DATASET.

	Covid	Non-covid	Total
Samples	53	78	131

At the time of admission to the hospital, the general condition of the patients can give information about the course of the disease. The increase in the respiratory rate per minute, which is normally 16-20, in these patients indicates that the disease is more affected[19].

Some blood values are measured in order to follow the course of the disease and to apply appropriate treatment. Leukocytes are the cells of the blood that fight infection. Neutrophil eosinophils and lymphocytes, subgroups of leukocytes, can be seen in different numbers and rates depending on the cause and intensity of the infection. These values provide important information about the severity and source of the disease. Lymphocytes generally play an important role in fighting viral infections. In viral infections, the number of lymphocytes is expected to increase, but in corona virus infections, the number is usually decreased. This decrease in lymphocyte count seems to be specific to covid infection and is closely related to the severity of the disease. Values such as sedimentation crp (c-reactive protein) and ldh (lactate dehydrogenase) are values that indicate the intensity of microbial infection in pneumonia. An increase can be seen in covid pneumonia as well as in bacterial pneumonia[19].

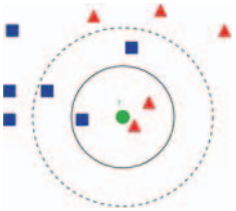
Platelets are cells that allow blood to clot. Fibrinogen and d-dimer are values indicating the tendency of the blood to clot. Studies have shown that Sars cov 2 virus is prone to clotting by disrupting the coagulation system, and if these values are high, blood thinners are used to prevent clotting in important organs such as the brain and lungs. In addition, increases in these values indicate that the disease is getting worse[19].

SARS-COV-2 virus is transmitted by inhalation and causes pneumonia. Unlike other causes of pneumonia, it may cause disorders by attaching to many organs and cells such as kidney, heart, liver, intestines, vessel wall, outside of the lung. ALT (alanine aminotransferase) AST (aspartate aminotransferase) are enzymes that increase in liver cell damage. Urea, creatinine are important values in showing the function of the kidneys. Troponin is an enzyme that increases when heart tissue is damaged. Increases in these values can provide information about the covid-related damage of the relevant organ. It can disrupt the mineral balance of the body by affecting all these organs. Therefore, blood sodium and potassium values are checked at the patient's arrival and follow-up[19].

Finally, we selected 24 parameters for training our model. Some of these are the respiratory rate per minute, neutrophil, eosinophils, lymphocytes, fibrinogen, d-dimer, ALT, AST, urea, creatinine, sodium and potassium values.

1. *k-Nearest-Neighbours Imputation*

As the initial dataset had missing data on some patients, we had to utilise an algorithm to make up for it. k-NN(k-Nearest Neighbours) is generally a quick and easy approach for finding relations in between samples. Each sample's missing data is imputed by finding the k closest neighbours of that particular point via Euclidean distance and taking their mean. One drawback for this algorithm is that it searches through all of the samples and is not suitable for larger datasets [16].



$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$$

Fig. 2. Figure and formula for k-Nearest-Neighbours implementation

In our case for 131 total samples, some features had more than 20 missing cells and we utilised kNN Imputation to make up new data which will not disturb the general behaviour of the model. Below graphs show the D-Dimer feature plotted over samples with its initial missing values and then imputed with kNN method when k=5 and k=10. We chose to continue our work with k=10 as it is generally accepted as a good value for most models and we observed it works better for our case as well. It can also be seen that the outliers do not affect the generated data.

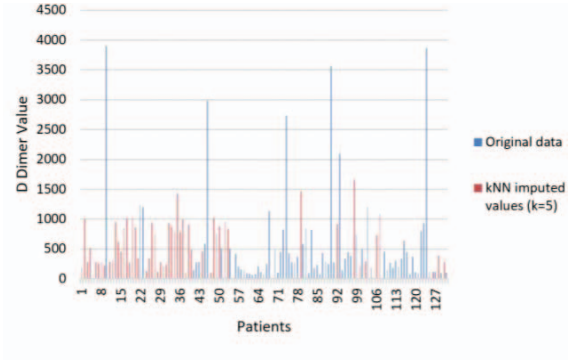


Fig. 3. kNN imputation of missing values

III. Proposed Models and Training

A. *Chest X-Ray Diagnosis Models*

After considering the models we can use, we have chosen resnet18 and squeezenet as our training models since better results had gotten from the previous research that used these models. For training these models we used transfer learning approach. By this way we fine tuned models that pre-trained on ImageNet dataset.

1. *Resnet18*

It is an 18-layered version of one of the most famous image recognizing algorithms that is being used nowadays called residual networks. As we mentioned previously this model brings a solution to the problem of vanishing gradients. In deep learning algorithms the weight update method is called gradient descent method. This method updates weights according to partial derivatives of the error function with respect to weights by using the chain rule. In some cases derivatives of the activation functions are

between -1 and 1. Therefore, multiplying derivatives in chain rule gradually decreases the rate. Sometimes even to the zero. This is called the vanishing gradients problem. The resnet algorithm is using a special method that provides shortcut connection from beginning layers to further layers in order to prevent this issue. This prevents gradients from exploid or vanishing and helps the learning process. These shortcuts are called identity mapping. The figure below indicates this process. Resnet50 and resnet101 basically use the same principle but have more layers. We chose resnet18 since the train process is more basic and fast[10].

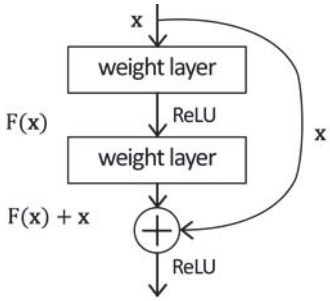


Fig. 4. A figure for showing ResNet18

2. Squeezenet

It is an algorithm that compresses the input parameters so that model size is smaller. The paper of the Squeezenet shows that even with the compression the accuracy is preserved and approximately the same as AlexNet. This accuracy is achieved with 50x fewer parameters and the total size of squeezenet is <0.5MB. Squeezenet uses three main strategies for it's compression. First they replace 3x3 convolution kernels with 1x1 ones. Then, they decrease the number of the input channels. Finally they use downsampling methods in order to get larger activation maps. These operations are made in the module called fire module. Below figure shows the fire module. Next figure shows the general representation of the squeezenet model[9].

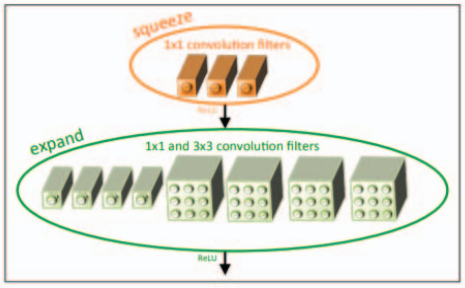


Fig. 5. A figure for showing SqueezeNet

3. Model Training and Parameters

In the output layer of our network we have used softmax regression which transforms outputs to probability values. Therefore in the training process we have utilized cross entropy loss function to calculate the error between target probability values and calculated probability values. To minimize this loss function we implemented Stochastic Gradient Descent Algorithm. This algorithm optimizes weight and bias values according to the partial derivatives of the loss function with respect to them. Amount of the change in optimization is determined with the learning rate. In previous studies that adopted resnet18 and squeezenet models, the best result has been achieved with a learning rate of 0.001. Therefore, we chose this value for our task.

$$S \longrightarrow \boxed{\text{Softmax}} \longrightarrow \boxed{\text{Cross -Entropy Loss}}$$

$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}}\ , CE = -\sum_i^C t_i \log f(s)_i$$

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

In the gradient descent method while converging to minimum loss, sometimes we encounter update values that deviate our way from the extremum point. To suppress this issue, we take account of the previous result of the update values of the weights and biases. To do this, in every step of the SGD, we take the previous gradient descent value and multiply with a coefficient called the momentum coefficient, then we add the result to the current gradient descent value to obtain the update value of the network parameters. We have taken the momentum coefficient as 0.9 because of the same reason as the learning rate.

$$\text{Repeat Until Convergence } \{$$

$$\nu \leftarrow \boxed{\eta * \nu_j} - \alpha \nabla_\omega \sum_1^m L_m(\omega)$$

$$\omega_j \leftarrow \nu_j + \omega_j$$

$$\}$$

We have trained our model for 100 epochs with the batch size of 20. We have trained our model by using CUDA technology with Nvidia GTX 1060 and 1050 graphical processing units. While training we have used cross validation technique to obtain training accuracy and training loss.

B. Blood Test Diagnosis

Considering the already numerical dataset and the low amount of required processing power, a multi layer neural network is utilised for this classification problem.

In the creation stage of the model, the number of hidden layers and epochs were set in order to inhibit outliers from affecting the learning process. The initial numbers of units for each hidden layer were determined as 8. Because of the same reason, the learning rate was chosen as 0.001. However, the model performed the overfitting issue even in this situation. To avoid the effect of overfitting, the number of units were updated to 8 and 4. For calculation of error between the target and predicted results, the previously explained cross entropy loss function was utilised. In addition to this, the Adam optimizer, which is an algorithm that builds predictions of low order moments, was used for updating weights in the back propagation stage. It is clear to see in the following graph, the Adam algorithm is more effective than other algorithms[11].

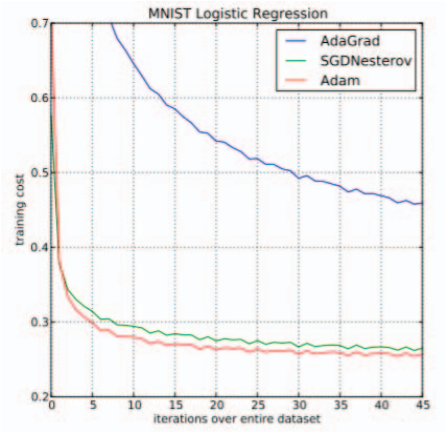


Fig. 6. A figure for showing Blood Test Graph

IV. Experimental Results

A. Chest X-Ray

We have divided this section of our research into two parts. First is a prediction between covid and no-finding X-Rays. Second one is prediction between covid and non-covid X-Rays. We have done this separation because in our non-covid X-Ray we have several other diseases which have X-Ray results similar to covid results that may confuse our network.

As evaluation criteria we have used sensitivity and specificity values since our dataset is unbalanced.

$$\text{Sensitiviy} = \frac{\text{Correct Prediction of Covid-19 Datas}}{\text{Total number of Covid-19 Datas}}$$

$$\text{Specificity} = \frac{\text{Correct Prediction of Non-Covid-19 Datas or No-Finding Datas}}{\text{Total number of Non-Covid-19 Datas or No-Finding Datas}}$$

We first obtained the histograms of probability values of the covid and non-covid samples. As can be seen from the figures our models are able to classify between covid and non-covid and the threshold value is around 0.2. Later we tried possible values between 0.1 and 0.3 to get the best accuracy.

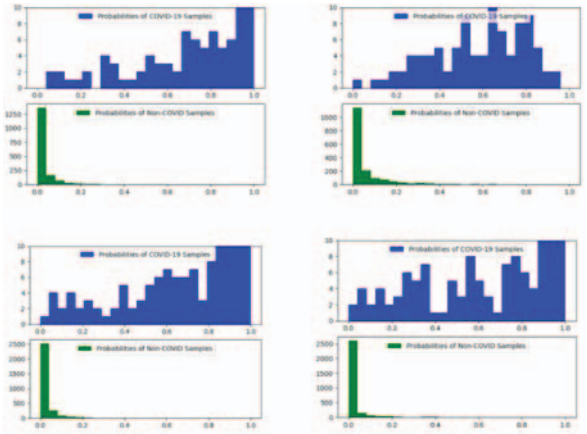


Fig. 7. Histograms

We obtained sensitivity and specificity for each of the models for 5 different threshold values. From the tables it is safe to say that the most optimum threshold value is 0.2.

These results can be shown visually by confusion matrices. For a threshold of 0.2 each of these matrices are combined in the table shown below.

TABLE IV. SCORE TABLES. THE 1st DATASET IS COVID vs NO FINDING, 2nd DATASET IS COVID VS NON-COVID

Model	Threshold	Sensitivity	Specificity	Time Elapsed(s)
ResNet18 on 1st dataset	0.1	%98	%91.9	66.50
	0.15	%95	%94.9	70.40
	0.2	%94	%96.5	66.16
	0.25	%92	%97.4	66.16
	0.3	%90	%98.1	66.71
ResNet18 on 2nd dataset	0.1	%96.6	%93.3	127.97
	0.15	%94.5	%95.6	126.30
	0.2	%91	%97.1	125.64
	0.25	%89	%97.9	129.60
	0.3	%86.9	%98.2	118.54
SqueezeNet on 1st dataset	0.1	%98	%81.6	72.53
	0.15	%98	%87.9	73.33
	0.2	%95	%91.4	74.37
	0.25	%93	%93.6	72.93
	0.3	%89	%94.8	73.80
SqueezeNet on 2nd dataset	0.1	%95.9	%93	122.92
	0.15	%93.8	%94.8	125.50
	0.2	%90.3	%96.2	126.42
	0.25	%88.3	%94.8	123.30
	0.3	%83.4	%97.4	125.01

As a result from these analyses, in covid vs no finding classification sensitivity is higher than covid vs non-covid one. This is a meaningful result since we have composed two different sources nearly with having covid number %50-%50. Therefore training with images from different sources makes classification harder but gives more reliable results. We have expected that specificity would increase along with the sensitivity too since no finding set does not include other diseases that may affect the model training. However specificity is smaller in the case of no finding. The reason can be that we have less data in the training set of the no finding model. Although both resnet18 and squeezenet are two most accurate methods to solve this problem, we can still see a slight advantage of resnet18 over squeezenet.

It is not exactly accurate to decide the performance of the model based on the sensitivity and specificity results as they differ with threshold values. Therefore we need to determine the accuracy independent from the threshold. There are several ways to handle this problem. In our project we implemented receiver operating curves (ROC). It indicates the rate of true positive samples with respect to the rate of false samples.

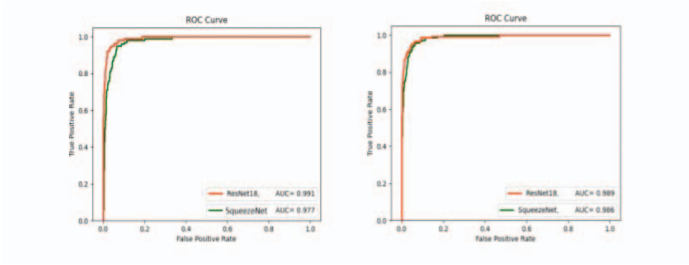


Fig. 8. Left plot representing the ROC of the 1st dataset and right plot representing the ROC of the 2nd dataset

From these results it can be inferred that resnet18 gives slightly better AUC which means better accuracy.

B. Blood Test

As our dataset includes only 131 samples, we needed to save as many samples as we can for the training phase so that the model could learn better but also keep enough so that we can correctly assess the accuracy of our model. We ended up using ¾ of our data for training and ¼ for testing

which were randomly distributed. Since our dataset consists mostly of non-covid cases, this unbalance also affects the output of the model if they were to split unevenly between the train and test sets. Thus we went with a model which was moderately separated.

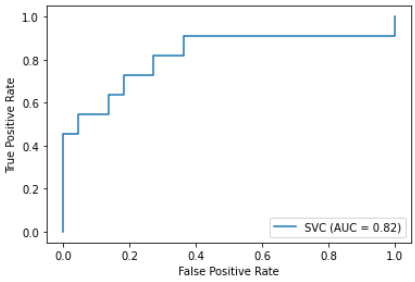


Fig. 9. ROC curve for blood test model

V. Conclusion

Considering the x-ray results, resnet18 performs slightly better than squeezenet. Generally, the scarcity of positive patients throughout the dataset has made the whole procedure unbalanced. Furthermore, even though the blood test dataset was not sufficient for a thorough research, we can deduce that it is a feasible method yet is not preferable over chest x-ray diagnosis.

In conclusion it can be seen that covid-19 diagnosis through various machine learning methodologies is possible and is theoretically applicable in modern medicine.

VI. Further Discussion

In our research, we classified whether the patient is covid or not by both using X-Ray data and blood test separately. Since the X-Ray data and blood tests are from different individuals, we could not combine them to obtain a total model which predicts with both the datasets. Moreover, in some of the previous research heatmaps or indications of possibly infected regions are produced[12, 16]. We did not implement that due to time limitations of our project. If these improvements are achieved in further studies, machine learning models can tremendously increase the information and the vision of health employees.

VII. Acknowledgement

Chest X-Ray and blood test result data has been taken from Canakkale Onsekiz Mart University with the permission number 2011-KAEK-27/2020-E.2000063714 of the ethics committee. During the research, any personal information was not used.

VIII. References

[1] Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2020). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Applied Intelligence, 1-11.

[2] A. Cozzi, S. Schiaffino, F. Arpaia, G. Della Pepa, S. Tritella, P. Bertolotti, et al. Chest x-ray in the COVID-19 pandemic: radiologists' real-world reader performance Eur J Radiol, 132 (2020), Article 109272, 10.1016/j.ejrad.2020.109272

[3] Alazab, M., Awajan, A., Mesleh, A., Abraham, A., Jatana, V., & Alhyari, S. (2020). COVID-19 prediction and detection using deep learning. International Journal of Computer Information Systems and Industrial Management Applications, 12, 168-181.

[4] Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.

[5] Asai, T. (2020). COVID-19: accurate interpretation of diagnostic tests—a statistical point of view.

[6] Cohen, J. P., Morrison, P., Dao, L., Roth, K., Duong, T. Q., & Ghassemi, M. (2020). Covid-19 image data collection: Prospective predictions are the future. arXiv preprint arXiv:2006.11988.

[7] Haykin, S. S. (2008). *Neural Networks and Learning Machines,third edition,* Pearson.

[8] 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).

[9] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. arXiv preprint arXiv:1602.07360.

[10] Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., ... & Ng, A. Y. (2019, July). Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 590-597).

[11] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[12] Minaee, S., Kafieh, R., Sonka, M., Yazdani, S., & Soufi, G. J. (2020). Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. Medical image analysis, 65, 101794.

[13] Narin, A., Kaya, C., & Pamuk, Z. (2020). Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003.10849.

[14] Maguolo, G., & Nanni, L. (2020). A critic evaluation of methods for covid-19 automatic detection from x-ray images. arXiv preprint arXiv:2004.12823.

[15] Monard, Maria-Carolina. (2002). A Study of K-Nearest Neighbour as an Imputation Method.

[16] 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), 1-12.

[17] Brinati, D., Campagner, A., Ferrari, D. *et al.* Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study. *J Med Syst* 44**,** 135 (2020)

[18] Abhirup Banerjee, Surajit Ray, Bart Vorselaars, Joanne Kitson, Michail Mamalakis, Simonne Weeks, Mark Baker, Louise S. Mackenzie, Use of Machine Learning and Artificial Intelligence to predict SARS-CoV-2 infection from Full Blood Counts in a population, International Immunopharmacology, Volume 86 (2020)

[19] AlJame, M., Ahmad, I., Imtiaz, A., & Mohammed, A. (2020). Ensemble learning model for diagnosing COVID-19 from routine blood tests. Informatics in Medicine Unlocked, 21, 100449.