COVID-19 detection on Chest X-Rays using Explainable Machine Learning

Subhajit Mondal, Subhankar Patra
M.Sc. Data and Knowledge Engineering
University of Magdeburg, Germany
{subhajit.mondal, subhankar.patra}@st.ovgu.de

Jalaj Vora, Shivam Singh
M.Sc. Digital Engineering
University of Magdeburg, Germany
{jalaj.vora, shivam.singh}@st.ovgu.de

Abstract—Starting from late December 2019, COVID-19 (SARS-CoV-2) has been declared a pandemic and a global emergency. Therefore, this disease has attracted utmost priority to be detected and hospitalize patients and treat them. In detecting the virus, RT-PCR kits are used which takes from hours to days for a diagnosis. This motivated the idea of building automated COVID-19 detector using artificial intelligence. There have been various deep learning based approaches built and achieve very high performance in diagnosis although, these architectures are un-interpretable. Therefore, this study reviews such deep learning based approches and focuses on building an ensemble architecture to diagnose COVID-19 by achieving comparable performance to the existing deep learning based COVID-19 detectors.

Index Terms—COVID-19 detection, Interpretable Machine Learning, Chest X-Rays, Radiology Level features, Explainibility

I. INTRODUCTION

From late December 2019, a novel corona-virus (SARS-CoV-2) has spread all around the globe originating from Wuhan district of China [1] [2]. As pf March 06, 2021 more than 116 million confirmed cases, and more than 2 million deaths were reported [1] around the world. Due to unavailability or difficult reach for immediate vaccination for novel COVID-19 disease, early diagnosis of the disease is highly critical. It provides the opportunity of immediate isolation of the suspected person and to decrease the chance of infection to healthy population. Reverse transcription polymerase chain reaction (RT-PCR) is used as main diagnosing method for COVID-19 [3], though it can be considered as a time-consuming test, as it takes typically hours or days to get the results and also, it suffers from false negative diagnosing. Chest radiography imaging (X-ray or computed tomography (CT) imaging) is used as a routine tool for pneumonia diagnosis and is easy to perform with fast diagnosis [4] [5] [7]. Chest CT has a high sensitivity for diagnosis of COVID-19 and X-ray images show visual indexes correlated with COVID-19 [6] [8].

The rapid use of CXRs were used by the radiology departments in Italy and the U.K. to sort non-COVID-19 patients with pneumonia to allocate hospital resources efficiently [9]. However, there are many common features between medical images of COVID-19 and pneumonia caused by other viral infections such as common flu (Influenzas-A). This similarity makes a differential diagnosis of COVID-19 cases by expert radiologists challenging [10], [11]. An explainable automated algorithm for classification of COVID-19 and non-COVID-19 CXR images can speed up the triage process of COVID-19 case detection and maximize the allocation of hospital resources.

Considering huge rate of suspected people and limited number of trained radiologists, machine learning methods for identification of such subtle abnormalities can assist the diagnosis procedure and increase the rate of early diagnosis with high accuracy. Machine Learning based solutions could be potentially powerful tools for solving such problems. Such an approach and intiative has already been shown by researchers especially using Deep learning based models more specifically using convolutional neural networks (CNN). These architectures have been shown to outperform the classical AI approaches in most of computer vision and and medical image analysis tasks in recent years, but this approach is considered "black box" due to it's complexity and un-explainibility [12]–[14]. It is tough to analyse for a person from medical domain or any individual why a system responded in the manner it did and raises the question of interpretability of the tool. The explainibility and relaibilty are the crucial factors in medical visual analytics. Therefore, we hypothesized that CXR images of COVID-19 patients can be reliably distinguished from other forms of pneumonia using an Machine Learning based classifier, along with radiological-level feature based explanation. We aimed to achieve similar or better performance compared to the existing Deep Learning Networks and also explain it's decisions.

This paper further discusses the motivation, ground truth and review of existing deep learning architectures in Section II. Section III proposes a machine-learning based ensemble architecture. Section IV describes implementation details and explores details of explainibility of the models in Section V. Section VI evaluates results and further concludes in Section VII.

II. BACKGROUND

The need to fasten the process of diagnosis or developing the diagnostic tool is more than ever for COVID-19 patients.

Typically, the results from RT-PCR kits take up-to approx. 6-8 hours to diagnose a patient being COVID-19 positive [14]. This motivated researchers the idea of using Chest Radiography Imaging especially, Chest X-Rays for diagnosis, as the Chest X-Rays are non-invasive tool to monitor progression of disease. Although, Chest CT Scan are considered high quality imaging but here experiment deals with Chest X-Rays due to it's high public availability.

A. Related Work

There are numerous experiments and studies built in order to apply Machine Learning and Deep Learning to assist diagnosis of COVID-19. Major amount of these studies are Deep Learning based architectures for COVID-19 classification. According to [16], there exists numerous studies which uses Statistical based feature extraction for COVID-19 detection. Though, these experiments also achieve comparable performance to the deep learning architecture, it has tough explainibility. These architectures are difficult to interpret in terms of their classification decision especially for any individual from medical background and radiologist. Also, study from [15]–[20] suggest that clinical and radiological-level features are crucial in identifying COVID-19 from Chest X-Rays.

Recently, many deep learning based architecture have been widely used for COVID-19 detection. Hemdan et al. [48] used deep learning models to diagnose COVID-19 in X-ray images and proposed a COVIDX-Net model comprising seven CNN models. Wang and Wong [46] proposed a deep model for COVID19 detection (COVID-Net), which obtained 92.4% accuracy in classifying normal, non-COVID pneumonia, and COVID-19 classes. Ioannis et al. [49] developed the deep learning model using 224 confirmed COVID-19 images. Their model achieved 98.75% and 93.48 % success rates for two and three classes, respectively. Narin et al. [37] achieved a 98% COVID-19 detection accuracy using chest X-ray images coupled with the ResNet50 model. Sethy and Behera [39] classified the features obtained from various convolutional neural network (CNN) models with support vector machine (SVM) classifier using X-ray images. Their study states that the ResNet50 model with SVM classifier provided the best performance. Finally, there are also several recent studies on COVID-19 detection that employed various deep learning models with CT images [52]–[57].

In this study, an ensemble framework is proposed for the diagnosis of COVID-19. The proposed framework has CheXNet architecture as feature extractor detecting 14 radiological features from raw Chest X-Ray images. These features are forwarded to the interpretable model which returns the diagnosis with feature-importance based explanation.

B. Choosen Models

We chose five simpler models namely; k-Nearest Neighbours, Support Vector Machine with Radial Basis Function kernel, Logistic Regression, Random Forest & Decision Trees and Multilayer Perceptron. Out of these five, we chose kNN, SVM with RBF, Logistic Regression and Decision Trees as interpretable simpler classical machine learning models and we considered Random Forest and MLP as Non-Interpreable Models or "Black Box" models. We tried to use "Black Box" model as our internal baseline to be trained on respective 14 features and compared it with the simpler interpretable models.

1) k-Nearest Neighbour: KNN is an algorithm that is considered both non-parametric and an example of lazy learning. kNN is a case-based learning method, which keeps all the training data for classification.However, to apply kNN we need to choose an appropriate value for k, and the success of classification is very much dependent on this value. In a sense, the kNN method is biased by k. There are many ways of choosing the k value, but a simple one is to run the algorithm many times with different k values and choose the one with the best performance.

2) SVM with RBF-kernel: A support vector machine constructs a hyper-plane or set of hyper-planes in a high or infinite dimensional space, which can be used for classification. Intuitively, a good separation is achieved by the hyper-plane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

RBF kernel is a function whose value depends on the distance from the origin or from some point. When training an SVM with the Radial Basis Function (RBF) kernel, two parameters must be considered: C and gamma. The parameter C, common to all SVM kernels, trades off misclassification of training examples against simplicity of the decision surface. A low C makes the decision surface smooth, while a high C aims at classifying all training examples correctly. gamma defines how much influence a single training example has. The larger gamma is, the closer other examples must be to be affected.

3) Logistic Regression: Another machine learning classification algorithm that is used to predict the probability of a categorical dependent variable. In logistic regression, the dependent variable is a binary variable that contains data coded as 1 (yes, success, etc.) or 0 (no, failure, etc.). Binary logistic regression requires the dependent variable to be binary. For a binary regression, the factor level 1 of the dependent variable should represent the desired outcome. Only the meaningful variables should be included. The independent variables should be independent of each other. That is, the model should have little or no multicollinearity. The independent variables are linearly related to the log odds. Logistic regression requires quite large sample sizes.

4) Random Forest: RF algorithm is one of the best algorithms for classification. RF is able for classifying large data with accuracy. It is a learning method in which number of decision trees are constructed at the time of training and outputs of the modal predicted by the individual trees. RF act as a tree predictors where every tree depends on the random vector values. The basic concept behind this is that a group of "weak learners" may come together to build a "strong learner". RF classifier is an ensemble method that trains

several decision trees in parallel with bootstrapping followed by aggregation, jointly referred as bagging. Bootstrapping indicates that several individual decision trees are trained in parallel on various subsets of the training dataset using different subsets of available features. Bootstrapping ensures that each individual decision tree in the random forest is unique, which reduces the overall variance of the RF classifier. For the final decision, RF classifier aggregates the decisions of individual trees; consequently, RF classifier exhibits good generalization. RF classifier tends to outperform most other classification methods in terms of accuracy without issues of overfitting.

5) Decision Trees: Decision tree builds classification or regression models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches, Leaf node represents a classification or decision. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data.

6) MultiLayer Perceptron: MLPs are suitable for classification prediction problems where inputs are assigned a class or label. They are comprised of one or more layers of neurons. Data is fed to the input layer, there may be one or more hidden layers providing levels of abstraction, and predictions are made on the output layer, also called the softmax layer. The power of neural networks comes from their ability to learn the representation in your training data and how to best relate it to the output variable that you want to predict. In this sense neural networks learn a mapping. Mathematically, they are capable of learning any mapping function and have been proven to be a universal approximation algorithm

7) Convolutional Neural Network: Convolutional Neural Networks, or CNNs, were designed to map image data to an output variable. CNN expect and preserve the spatial relationship between pixels by learning internal feature representations using small squares of input data. Feature are learned and used across the whole image, allowing for the objects in the images to be shifted or translated in the scene and still detectable by the network. It is this reason why the network is so useful for object recognition in photographs, picking out digits, faces, objects and so on with varying orientation. They have proven so effective that they are the go-to method for any type of prediction problem involving image data as an input.

III. PROTOTYPE DESIGN

From our project, the research questions concentrate on the degree of explainability resulted by simpler classifier when used on 14 features extracted from CheXNet architecture.

A. Research Questions

1) Does the simpler Machine learning achieves results as good as deep learning based classifiers while being better explainable with regards to input features ?

2) Are the SHAP features extracted from Machine Learning models are interpretable by Radiologist for COVID-19 detection ?

B. Structure / Design / Pipeline

The state of art for image classification deep learning states that image is classified based upon the feature learned by the model against the class.

The explainability for the deep learning model for the CXR images of COVID affected patient could reveal the parts/patches from the image that are highly influence the prediction task. However, this visual explanation could not explain biological features responsible for the prediction. We experimented with simpler machine learning models like SVM, kNN, Logistic Regression, Random Forest, Decision Tree. For each image in the dataset we have generated the value of 14 radiological features that are Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis, Pleural Thickening Hernia. As it is a supervised learning task, the simpler machine learning models learn to classify into COVID positive or COVID negative based upon these 14 radiological features. Further, we extracted the SHAP values from the SHAP explainer that shows the contribution by each feature responsible for the decision making process. Further, in order to experiment with SHAP values we have normalised accuracies of all the models which will be used to get the consolidated SHAP values for each model using Weighted Average method. We tried to find out what relation could be the established between local and global interpretability by clustering SHAP values.
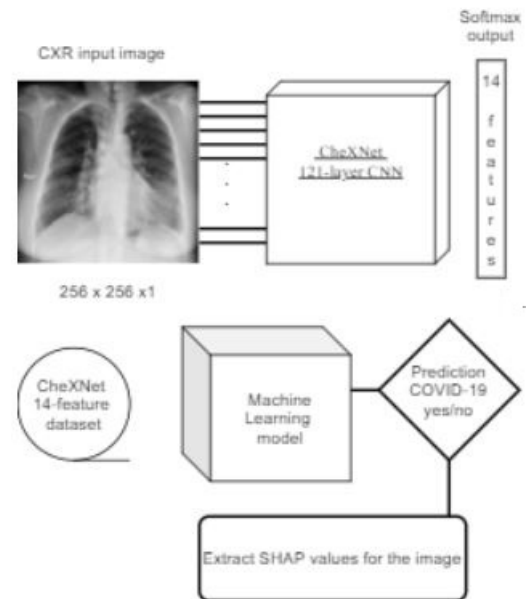

Fig. 1. Project Structure

C. Dataset Generation

ChexNet is a deep learning algorithm that can detect and localize 14 kinds of pathology features from chest X-ray images. As described in the paper, a 121-layer densely connected convolutional neural network is trained on ChestX-ray14 dataset, which contains 112,120 frontal view X-ray images from 30,805 unique patients. The result is so good that it surpasses the performance of practicing radiologists.

We used pre-trained CheXNet model for generating the 14 pathological features for each images in the dataset with label as covid positive or negative. These 14 pathological feature value of a image serves as a data point. Therefore, the datset consists of columns represented by number of features and label while rows represents number of images.

D. Evaluation Metrics

Choosing an appropriate evaluation metric is challenging generally in applied machine learning, but is particularly difficult for imbalanced classification problems. As most of the standard metrics that are widely used assume a balanced class distribution, and because typically not all classes, and therefore, not all prediction errors, are equal for imbalanced classification. Therefore, we tried to evaluate based upon the correct classification given by model. We choose Accuracy, Receiver Operating Characteristic and F1 score for evaluating performance of each model.

Accuracy: It is the ratio of number of correct predictions to the total number of input samples.

ROC curve:It is a diagnostic plot for summarizing the behavior of a model by calculating the false positive rate and true positive rate for a set of predictions by the model under different thresholds.

F1 score: F1 Score is the Harmonic Mean between precision and recall. The range for F1 Score is [0, 1]. It tells you how precise your classifier is (how many instances it classifies correctly), as well as how robust it is (it does not miss a significant number of instances).

Further, we considered a benchmark dataset to evaluate performance of the models for data coming from different distribution. The database was developed by [59] using images from various sources. The database is constantly upgraded, as of now, the database content comprise of around "201" COVID-19 positive X-Ray images. There were no images, which represent X-Ray of normal lungs, hence we have used thousand healthy X-Ray images from our original database as negative class, which we have used to feed our original models.

IV. IMPLEMENTATION

A. Datasets

Our Dataset consists of 313 Positive COVID CXR and 1000 Negative CXR collected from four different sources to make our version of the dataset to work upon. This includes COVIDx dataset from [59], Kaggle CXR Pneumonia dataset by Paul Mooney [60], CXR images of adult subjects from the RSNA Pneumonia Detection Challenge [61], original and augmented

versions of COVID-19 examples from [62]. We split the data set in ratio 0.95:0.05 and trained the models with 1247 data points and 66 test points.

Data-set	COVID	Non-COVID
	313	1000

TABLE I
OVERVIEW OF GROUND TRUTH

According to [63]–[67] CT-Scan data would be gold-standard for us and also portray satisfying results when evaluated in terms of Accuracy and F1-Score. However, due to CT Scan being available in very less quantity publicly, we would like to use Chest X-rays as our dataset. Though, it won't be that compatible in terms of quality in regards with CT-Scans but [68] suggests CXR to be sufficient and comparable to CT-Scans in order to diagnose COVID-19 patients.

Real world datasets are mostly imbalanced. The COVID-19 CXR (Chest X-Rays) dataset we used has around 1000 to 313 negative to positive ratio. That there are nearly three times more negative cases than that of positive. The classification algorithms in this case tends to favor the majority class. The distribution of the classes in the dataset in reality refers to the actual class distribution of the covid affected cases. Hence, the learning task in hand is a imbalanced learning task because there are more people with healthy lungs than that of with COVID-19.

B. Models

The choice of the simpler machine learning models that can explains the model's decision for the classification and should results in minimum false positive and false negative.

1) kNN: Model finds 9-nearest neighbour for each sample, we arrived to k=9 from the Elbow-plot and Silhouette score for the dataset.

2) SVM: Support Vector Machine with Radial Basis Function(RBF) along with 5-fold cross validation internally as probability estimate is enabled.

3) Logistic Regression: Logistic Regression with internal cross-validation(only 2) and L2 Regularizer.

4) Decision Tree: Using 'gini' function to measure the quality of a split with minimum 2-samples required to split an internal node

5) Random Forest: We used similar 'gini' function to measure the quality of a split with minimum 2-samples required to split an internal node for a tree. The forest consist of 100 trees.

6) MLP: The network consits of two hidden layers with Adam optimizer, L2 Regularizer and learning rate is 0.001.

V. EXPLAINAINIBILITY

The Shapley values obtained using the Shap algorithm for each axis, gives the importance for each axis of an instance. Average of the absolute values of the shapley values (per axis) thus obtained gives us the global shapley values per axis. And thus, we have importance of axis per instance and globally. The experiment we have conducted tried to answer,

if and how can we compare global shapley values, with ones obtained for single instances. For some data, we have the following summary plot of shapley values which gives the global importance and further for some single instance, we have the following barplot
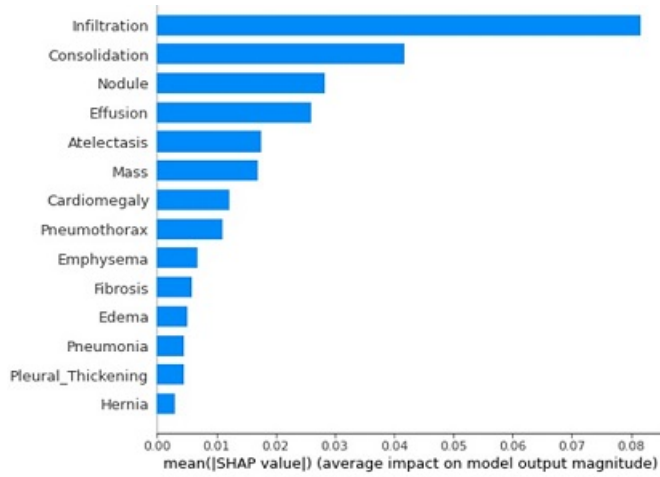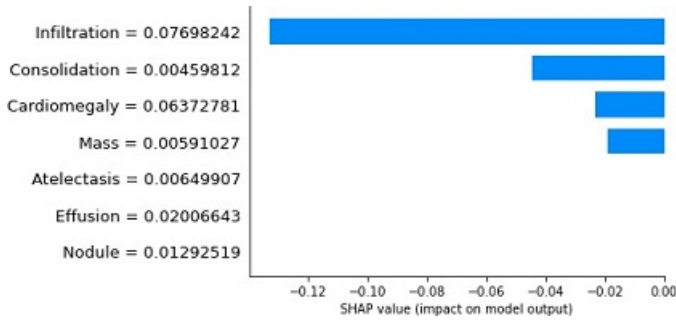

Fig. 2. Global Shap values for kNN

Fig. 3. Shap values for kNN of an instance

Both combined in a single plot we have the following when axis are sorted by importance of global features
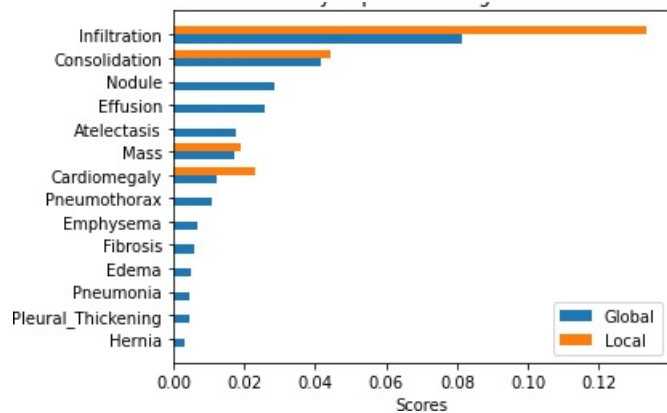

Fig. 4. Feature importance

We can see for cardiomegaly and mass axis, the single instance provides high importance on axis that have low importance in case of global plot. Also, the ranks induced by the shapley values of the cardiomegaly and mass in case of single instance is opposite to what it is in global case. We put up a hypothesis that the importance surface induced by the individual instances, is a non-concave surface, and thus have multiple peaks or local maxima, where the importance of axis are different from one peak to another. Thus, the motivation is that, we might get a better interpretation when comparing one single instance with the nearest maxima induced axis importance.

To achieve this, we cluster the data. Now for each cluster we find the global inter-cluster importance. Now we only compare one single instance and the inter-cluster global importance of axis of the cluster the single instance belongs to. For this case, the single instance falls to cluster 2 and thus we compare the shapley values of the instance with the inter-cluster global importance of cluster 2. The following is the plot when axis are sorted by importance induced by the inter cluster importance of axis.
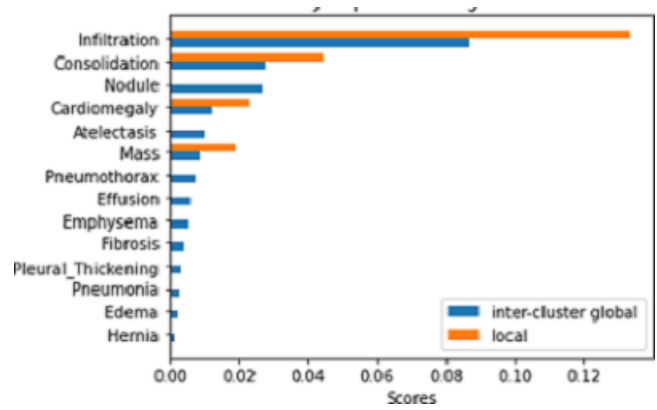

Fig. 5. Feature importance

We can see now that, the axis of a single instance has nearly the same proportional weights to axis it gives importance to as that of inter-cluster global importance. And thus, we can say, we get better comparison of a single instance to inter-cluster global importance of the cluster the instance belongs to rather than the actual global importance. However, again this behavior is controlled by the distance of the instance from the cluster center. Experiment shows, as the distance from the centroid of the cluster increases for a single instance, the results starts diverging from the above said pattern. And thus, the comparison of a single instance to inter-cluster global importance of the cluster the instance belongs to, is no worse than comparing to the actual global importance.

In the above example it can be seen that both cases fibrosis, which is 9th rank in both inter-cluster global importance and actual global importance, but given higher importance, by local instance. However, if we are to compare the other axis plots, we see the single instance follows the axis importance induced

by the inter-cluster global importance better rather than the actual global importance.
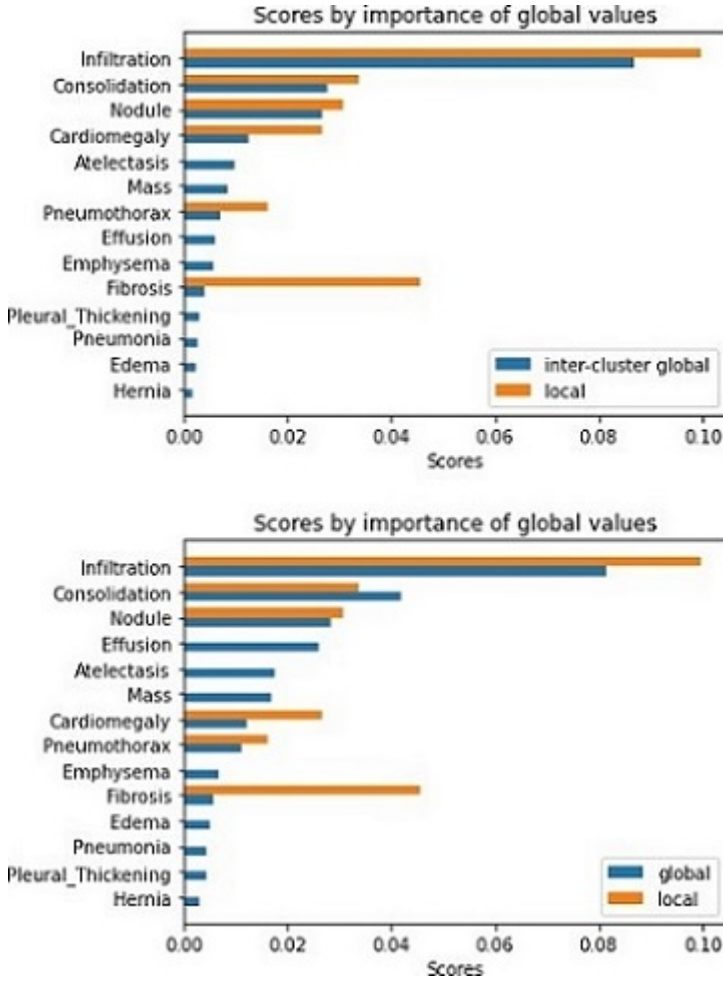

Fig. 6. Feature importance

So it can be concluded that, the inter-cluster global importance of the clusters provide no worse comparison to the actual global importance. Also to note is, the actual global importance of axis, can be better compared with the inter-cluster global importance than the individual instances.

VI. RESULTS AND DISCUSSION

Here II outlines the score on test set of the 5 chosen interpretable models. Though our baseline model MLP trained on the ground truth achieves 98.7% of accuracy, the simpler models do not achieve the similar outcome as our baseline.

Measures		Models				
		kNN	SVM	LR	DT	RF
Original dataset	Accuracy	86%	80.30%	80%	74%	81.81%
	AuC	0.73	0.78	0.71	0.67	0.87
	F1score	0.66	0.65	0.63	0.64	0.69

TABLE II
EVALUATION OF THE MODELS

VII. CONCLUSION

In this study, we aimed to build a system which can be a visual aid to the medical personal dealing with diagnosis of COVID-19. We hypothesis to achieve a comparable performance to the deep learning based architecture and our internal baseline model by using simpler and more interpretable 5 machine learning models. We conducted the experiment where we also came up with the explaination of the classification using SHAP graphs. The graphs explains the decision based on features used by the model stating values to be higher the greater.

The section VI shows that the experiment didn't out perform the baseline MLP but the models were more explicit and interpretable. Therefore, we would like to study more introspection to the models to better tune them to achieve higher performance than the internal baseline MLP and the related deep learning based architectures.

Future work can be done by fine tuning the ensemble framework and models and also by comparing with existing deep learning architectures. Also, a user study can be conducted with the help of radiologists to evaluate more objectively.

REFERENCES

[1] World Health Organization. Director-General's opening remarks at the media briefing on COVID-19 - March 11, 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.
[2] Wang, Chen, et al. "A novel coronavirus outbreak of global health concern." The lancet 395.10223 (2020): 470-473.
[3] Cohen, Jon. "Wuhan seafood market may not be source of novel virus spreading globally." Science 10 (2020).
[4] Ai, Tao, et al. "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases." Radiology 296.2 (2020): E32-E40.
[5] Long, Chunqin, et al. "Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT?." European journal of radiology 126 (2020): 108961.
[6] Zu, Zi Yue, et al. "Coronavirus disease 2019 (COVID-19): a perspective from China." Radiology 296.2 (2020): E15-E25.
[7] Lee, Elaine YP, Ming-Yen Ng, and Pek-Lan Khong. "COVID-19 pneumonia: what has CT taught us?." The Lancet Infectious Diseases 20.4 (2020): 384-385.
[8] Rubin, Geoffrey D., et al. "The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society." Chest 158.1 (2020): 106-116.
[9] Castiglioni, Isabella, et al. "Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy, Italy." medRxiv (2020).
[10] Neuman, Mark I., et al. "Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children." Journal of hospital medicine 7.4 (2012): 294-298.
[11] Ng, Ming-Yen, et al. "Imaging profile of the COVID-19 infection: radiologic findings and literature review." Radiology: Cardiothoracic Imaging 2.1 (2020): e200034.
[12] Ahsan, Md Manjurul, et al. "Study of different deep learning approach with explainable ai for screening patients with COVID-19 symptoms: Using ct scan and chest x-ray image dataset." arXiv preprint arXiv:2007.12525 (2020).

[13] Akl, Elie A., et al. "Use of chest imaging in the diagnosis and management of COVID-19: a WHO rapid advice guide." Radiology 298.2 (2021): E63-E69.

[14] World Health Organization. (2020). Use of chest imaging in COVID-19: a rapid advice guide, 11 June 2020. World Health Organization. https://apps.who.int/iris/handle/10665/332336. Licencia: CC BY-NC-SA 3.0 IGO

[15] Won, Joungha, et al. "Development of a Laboratory-safe and Low-cost Detection Protocol for SARS-CoV-2 of the Coronavirus Disease 2019 (COVID-19)." Experimental neurobiology 29.2 (2020): 107.

[16] Zhang, Zhenwei, and Ervin Sejdić. "Radiological images and machine learning: trends, perspectives, and prospects." Computers in biology and medicine 108 (2019): 354-370.

[17] Sánchez-Oro, Raquel, Julio Torres Nuez, and Gloria Martínez-Sanz. "Radiological findings for diagnosis of SARS-CoV-2 pneumonia (COVID-19)." Medicina clinica (English ed.) (2020).

[18] Zhang, Ran, et al. "Diagnosis of coronavirus disease 2019 pneumonia by using chest radiography: Value of artificial intelligence." Radiology 298.2 (2021): E88-E97.

[19] Cleverley, Joanne, James Piper, and Melvyn M. Jones. "The role of chest radiography in confirming covid-19 pneumonia." bmj 370 (2020).

[20] Cozzi, Diletta, et al. "Chest X-ray in new Coronavirus Disease 2019 (COVID-19) infection: findings and correlation with clinical outcome." La radiologia medica 125 (2020): 730-737.

[21] Shi, Heshui, et al. "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study." The Lancet infectious diseases 20.4 (2020): 425-434.

[22] Huang, Chaolin, et al. "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China." The lancet 395.10223 (2020): 497-506.

[23] Zheng, Qiuting, et al. "Clinical and radiological features of novel coronavirus pneumonia." Journal of X-ray Science and Technology 28.3 (2020): 391-404.

[24] Brunese, Luca, et al. "Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays." Computer Methods and Programs in Biomedicine 196 (2020): 105608.

[25] Toba, Shuhei, et al. "Prediction of Pulmonary to Systemic Flow Ratio in Patients With Congenital Heart Disease Using Deep Learning–Based Analysis of Chest Radiographs." JAMA cardiology 5.4 (2020): 449-457.

[26] Toussie, Danielle, et al. "Clinical and chest radiography features determine patient outcomes in young and middle-aged adults with COVID-19." Radiology 297.1 (2020): E197-E206.

[27] Durrani, Misbah, Ume Kalsoom Inam ul Haq, and Anum Yousaf. "Chest X-rays findings in COVID 19 patients at a University Teaching Hospital-A descriptive study." Pakistan Journal of Medical Sciences 36.COVID19-S4 (2020): S22.

[28] Apostolopoulos, Ioannis D., Sokratis I. Aznaouridis, and Mpesiana A. Tzani. "Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases." Journal of Medical and Biological Engineering 40 (2020): 462-469.

[29] Yasin, Rabab, and Walaa Gouda. "Chest X-ray findings monitoring COVID-19 disease course and severity." Egyptian Journal of Radiology and Nuclear Medicine 51.1 (2020): 1-18.

[30] Rajpurkar, Pranav, et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." arXiv preprint arXiv:1711.05225 (2017).

[31] Cohen, Joseph Paul, et al. "On the limits of cross-domain generalization in automated X-ray prediction." Medical Imaging with Deep Learning. PMLR, 2020.

[32] Cohen, Joseph Paul, et al. "Predicting covid-19 pneumonia severity on chest x-ray with deep learning." Cureus 12.7 (2020).

[33] Han, Yan, et al. "Pneumonia Detection on Chest X-ray using Radiomic Features and Contrastive Learning." arXiv preprint arXiv:2101.04269 (2021).

[34] Tsai, Min-Jen, and Yu-Han Tao. "Machine Learning Based Common Radiologist-Level Pneumonia Detection on Chest X-rays." 2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS). IEEE, 2019.

[35] Kumar, Gaurav, and Pradeep Kumar Bhatia. "A detailed review of feature extraction in image processing systems." 2014 Fourth international conference on advanced computing & communication technologies. IEEE, 2014.

[36] Hong, Zi-Quan. "Algebraic feature extraction of image for recognition." Pattern recognition 24.3 (1991): 211-219.

[37] Santosh, K. C. "AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on multitudinal/multimodal data." Journal of medical systems 44.5 (2020): 1-5.

[38] Sahlol, Ahmed T., et al. "COVID-19 image classification using deep features and fractional-order marine predators algorithm." Scientific Reports 10.1 (2020): 1-15.

[39] Sethy, Prabira Kumar, et al. "Detection of coronavirus disease (COVID-19) based on deep features and support vector machine." (2020).

[40] Perumal, Varalakshmi, Vasumathi Narayanan, and Sakthi Jaya Sundar Rajasekar. "Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features." Applied Intelligence 51.1 (2021): 341-358.

[41] Pereira, Rodolfo M., et al. "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios." Computer Methods and Programs in Biomedicine 194 (2020): 105532.

[42] Khuzani, Abolfazl Zargari, Morteza Heidari, and S. Ali Shariati. "COVID-Classifier: An automated machine learning model to assist in the diagnosis of COVID-19 infection in chest x-ray images." medRxiv (2020).

[43] Elaziz, Mohamed Abd, et al. "New machine learning method for image-based diagnosis of COVID-19." Plos one 15.6 (2020): e0235187.

[44] Tsiknakis, Nikos, et al. "Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays." Experimental and Therapeutic Medicine 20.2 (2020): 727-735.

[45] Zokaeinikoo, Maryam, et al. "Aidcov: An interpretable artificial intelligence model for detection of covid-19 from chest radiography images." medRxiv (2020).

[46] 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.

[47] Khan, Asif Iqbal, Junaid Latief Shah, and Mohammad Mudasir Bhat. "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images." Computer Methods and Programs in Biomedicine 196 (2020): 105581.

[48] 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).

[49] Apostolopoulos, Ioannis D., and Tzani A. Mpesiana. "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks." Physical and Engineering Sciences in Medicine 43.2 (2020): 635-640.

[50] Narin, Ali, Ceren Kaya, and Ziynet Pamuk. "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks." arXiv preprint arXiv:2003.10849 (2020).

[51] DeepCOVIDExplainer: Explainable COVID-19 Diagnosis from Chest X-ray Images, Karim, Md Rezaul and Döhmen, Till and Rebholz-Schuhmann, Dietrich and Decker, Stefan and Cochez, Michael and Beyan, Oya, IEEE International Conference on Bioinformatics and Biomedicine (BIBM'2020), IEEE, 2020

[52] Hasan, Ali M., et al. "Classification of covid-19 coronavirus, pneumonia and healthy lungs in ct scans using q-deformed entropy and deep learning features." Entropy 22.5 (2020): 517.

[53] Barstugan, Mucahid, Umut Ozkaya, and Saban Ozturk. "Coronavirus (covid-19) classification using ct images by machine learning methods." arXiv preprint arXiv:2003.09424 (2020).

[54] 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.

[55] Wu, Yu-Huan, et al. "Jcs: An explainable covid-19 diagnosis system by joint classification and segmentation." IEEE Transactions on Image Processing (2021).

[56] Kassani, Sara Hosseinzadeh, et al. "Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: A machine learning-based approach." arXiv preprint arXiv:2004.10641 (2020).

[57] Chen, Yuanfang, et al. "An interpretable machine learning framework for accurate severe vs non-severe covid-19 clinical type classification." Available at SSRN 3638427 (2020).

[58] Chatterjee, Soumick, et al. "Exploration of interpretability techniques for deep covid-19 classification using chest x-ray images." arXiv preprint arXiv:2006.02570 (2020).

[59] https://medium.com/capital-one-tech/k-nearest-neighbors-knn-algorithm-for-machine-learning-e883219c8f26

[60] Gongde Guo1, Hui Wang , David Bell , Yaxin Bi, and Kieran Greer School of Computing and Mathematics, University of Ulster Newtownabbey, BT37 0QB, Northern Ireland, UK G.Guo,H.Wang,Krc.Greer@ulst.ac.uk 2 School of Computer Science, Queen's University Belfast, Belfast, BT7 1NN, UK DA.Bell,Y.Bi@qub.ac.uk (kNN)

[61] https://scikit-learn.org/stable/modules/svm.html

[62] https://www.sciencedirect.com/topics/engineering/random-forest

[63] https://machinelearningmastery.com/neural-networks-crash-course/

[64] https://machinelearningmastery.com/crash-course-convolutional-neural-networks/

[65] Chest X-Ray Images (Pneumonia) by PaulcMooney https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

[66] RSNA Pneumonia Detection Challenge https://www.kaggle.com/c/rsna-pneumonia-detection-challenge

[67] COVID-19 Image Data Collection: Prospective Predictions Are the Future Joseph Paul Cohen and Paul Morrison and Lan Dao and Karsten Roth and Tim Q Duong and Marzyeh Ghassemi arXiv:2006.11988, https://github.com/ieee8023/covid-chestxray-dataset, 2020

[68] Zhao, Y. Zhang, X. He, P. Xie, COVID-ct-dataset: A ct scan dataset about covid-19, ArXiv. abs/2003.13865 (2020).

[69] L. Wang, A. Wong, COVID-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, arXiv Preprint arXiv:2003.09871. (2020).

[70] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, others, Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct, Radiology. (2020) 200905.

[71] D. Singh, V. Kumar, M. Kaur, Classification of covid-19 patients from chest ct images using multi-objective differential evolution–based convolutional neural networks, European Journal of Clinical Microbiology & Infectious Diseases. (2020) 1–11.

[72] M. Karim, T. Döhmen, D. Rebholz-Schuhmann, S. Decker, M. Cochez, O. Beyan, others, Deepcovidexplainer: Explainable covid-19 predictions based on chest x-ray images, arXiv Preprint arXiv:2004.04582. (2020).

[73] D.S. Kermany, M. Goldbaum, W. Cai, C.C.S. Valentim, H. Liang, S.L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, J. Dong, M.K. Prasadha, J. Pei, M.Y.L. Ting, J. Zhu, C. Li, S. Hewett, J. Dong, I. Ziyar, A. Shi, R. Zhang, L. Zheng, R. Hou, W. Shi, X. Fu, Y. Duan, V.A.N. Huu, C. Wen, E.D. Zhang, C.L. Zhang, O. Li, X. Wang, M.A. Singer, X. Sun, J. Xu, A. Tafreshi, M.A. Lewis, H. Xia, K. Zhang, Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell. 172 (2018) 1122–1131.e9. https://doi.org/https://doi.org/10.1016/j.cell.2018.02.010.