

COVID-19 Identification using Machine Learning Classifiers with GLCM Features of Chest X-ray Images

Sudeep D. Thepade* and Hrishikesh Jha

Department of Computer Engineering, Pimpri Chinchwad College of Engineering, SPPU, Maharashtra, India

(*Corresponding author's e-mail: sudeepthepade@gmail.com)

Received: 20 August 2020, Revised: 10 June 2021, Accepted: 20 June 2021

Abstract

COVID-19 is an ongoing pandemic, and is also known by the name coronavirus. It was originally discovered in Wuhan, China, in December, 2019. Since then, it has been increasing rapidly worldwide. Since it has been increasing at such a rapid pace, testing equipment has limited availability. Also, this disease spreads very quickly, so it is better if it is detected earlier, in order so that it can be stopped from spreading. Therefore, the importance of early detection has increased; however, because of the shortage of testing sets, it is a necessity to develop an automated system that can detect whether the COVID-19 disease is present in a person or not as early as possible. Therefore, in this work, to extract features from X-ray images of the chest, we have made use of the Gray Level Co-occurrence Matrix (GLCM). After extracting these features for the classification of the images, we used different machine learning models, and an ensemble of machine learning models, to classify X-ray images of the chest as COVID-19, Normal, Pneumonia-bac, or Pneumonia-vir. Considering the average of performance metrics, the ensemble of Random Forest-MLP gave the best result among the variations.

Keywords: GLCM, COVID-19, X-ray, Pneumonia-bac, Pneumonia-vir, Machine learning

Introduction

COVID-19, or the novel coronavirus pandemic, seemed to appear in Wuhan, China [1], and since then has become a serious health problem across the world [2]. It is instigated by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus is thought to have a zoonotic origin [3], which means that it has been caused by a pathogen that has jumped from an animal to a human. This infected human then transmits the infectious agent to other persons, spreading it.

It can spread from one infected person to another. Primarily, it spreads through small droplets from the nose or mouth of a COVID-19 infected person when they sneeze, cough, or speak [4]. A person can also become affected by breathing in these droplets. If the person is unaware of the infection of another person, then they might think that they are safe but, in reality, they are spreading the virus even more. So, it becomes important that the infection is detected as early as possible, so that the spreading of the virus can be stopped.

Some of the symptoms of COVID-19 include fatigue, alteration in the sense of smell or taste, shortness of breath, sore throat, headache, and vomiting [5]. As the symptoms of coronavirus seem quite similar to the symptoms of viral flu, it becomes difficult to distinguish between them without proper testing in healthcare centers.

Since any kind of vaccine or treatment has not yet been developed for novel coronavirus disease, it becomes important to detect the infected person in the early stages and isolate them from the healthy population. COVID-19 diagnosis after detecting the symptoms needs to be confirmed using RT-PCR test (Reverse Transcription-Polymerase Chain Reaction testing) of infected secretions or chest CT imaging [6]. According to a study, a chest CT scan or chest X-ray for diagnosis of COVID-19 disease gave high sensitivity results, which make it better to be used as a COVID-19 diagnosis in areas affected by this pandemic.

Since testing kits are being used a lot, there is a shortage, which will make automated detection of COVID-19 a necessity in times to come. Hence, this paper attempts to classify chest X-ray images into different classes, which are COVID-19, Normal, Pneumonia-bac, or Pneumonia-vir. Extraction of

features from chest X-ray images is done using Gray Level Co-occurrence Matrix (GLCM), which are then classified using several machine learning models and an ensemble of machine learning models.

The rest of the paper proceeds as follows. The next section covers the literature survey conducted for the paper. After that, we cover our proposed method. Next, we cover the experimentation environment in which the experiments were considered. In the section for results and discussion, we discuss the results of the experiment. Finally, we give concluding remarks on the paper.

Literature survey

COVID-19 has emerged as one of the deadliest diseases and is a dangerous pandemic for the whole world. It is affecting all the sectors in every corner of the world.

Since COVID-19 is still new, a lot of research on this topic remains to be done, and because of this, proper diagnostic testing for this virus are still not available. Similarly, as it is so new, medical professionals are not equipped with the proper training and knowledge to deal with this virus.

One of the possible ways of diagnosis is to obtain a chest X-ray of the patient and identify whether they have COVID-19 or not. However, this method is not fully reliable as, in a visible way, it is quite similar to lung infection in pneumonia. Also, because of its similarity with pneumonia in X-ray, it becomes difficult to distinguish from the resemblance between COVID-19 or viral pneumonia, and bacterial pneumonia. For pneumonia and COVID-19 detection in X-ray images, GLCM was used earlier in studies, as follow.

In paper [7], a computer-aided system for chest radiography, which identified bacterial and viral pneumonia, was given. In this method, lung regions are first identified and then classified in the pneumonia category. In the beginning, regions of the lung are extracted using a fully convolutional networks model, and then tested on the Montgomery County, Maryland, USA, dataset and Japanese Society of Radiological Technology database. Succeeding segmentation, in order to allow for the classification of target lung regions, a model of a deep convolutional neural network is used; then, these features are compared to manual features, and both of these features are fused. Then, it is finally given to the SVM classifier for binary classification.

Paper [8] classified lung disease using deep networks and compared the results with the results of GLCM-based classification. The dataset used here consisted of information on 81 infected and 15 normal patients as computed tomography slices of high resolution. Various networks of deep learning were used, namely, Alexnet, VGG16, VGG19, Res59 and Res101. Then, it was classified using 5 classifiers, and then Principal Component Analysis (PCA) was used, with increased classification accuracy.

In paper [9], the dataset consisted of 100 images of patients affected by COVID-19 and normal patients acquired from the GitHub repository and Kaggle repository. Three convolutional neural network models, InceptionV3, ResNet50, and Inception-ResNetV2, were used. Then, ROC and confusion matrices were used and evaluated using 5-fold cross-validation. It was observed that the highest classification performance was given by the ResNet50 model, with an accuracy of 98 %.

In [10], Arpan *et al.* used a dataset with X-ray images of the chest, classifying them into 4 classes, normal, bacterial pneumonia, viral pneumonia, and COVID-19. They used a pre-trained deep learning model, CheXNet, with a 121-layer Dense Convolutional Network (DenseNet) backbone, which was preceded by a fully connected layer. Using this, they achieved an accuracy of 87.2 %.

Using Deep Neural Networks will give a higher accuracy, but also requires a lot more data that is not readily available. So, here, we have used GLCM to extract textural features from the images and have then classified them using machine learning models and a majority voting-based ensemble of machine learning models.

Proposed methodology

The X-ray images are first obtained from the dataset, and then, by using Gray Level Co-occurrence Matrix (GLCM), their features are extracted. GLCM is a matrix that extracts textural features from images. These extracted features are then collected and made into a feature set. On this feature set, we train several machine learning models and variations of the ensemble of machine learning models. We have used a method of taking 1 class as positive and the other 3 as negative at 1 time to calculate performance. For example, to check the performance of the COVID-19 class, we took the COVID-19 class as positive, and the other 3 (Normal, Pneumonia-bac, and Pneumonia-vir) as negative. Similarly, we have done this for all 4 classes.

The flow diagram given in **Figure 1** demonstrates the working of the proposed COVID-19 detection system using GLCM features with machine learning models.

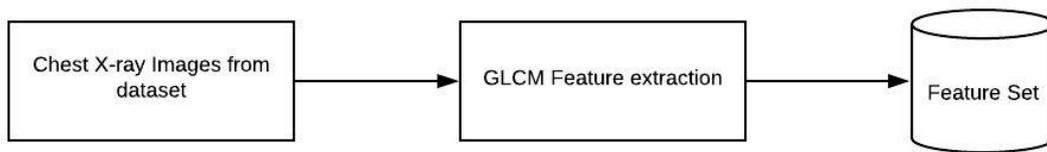


Figure 1 Phase 1 - Feature Extraction using GLCM for proposed COVID-19 detection.

These feature sets are then given to several machine learning algorithms, as shown in **Figure 2**. The top-performing algorithms are then selected to form ensembles using the concept of majority voting. These trained models then classify the image between the 4 classes, namely, COVID-19, Normal, Pneumonia-bac, and Pneumonia-vir.

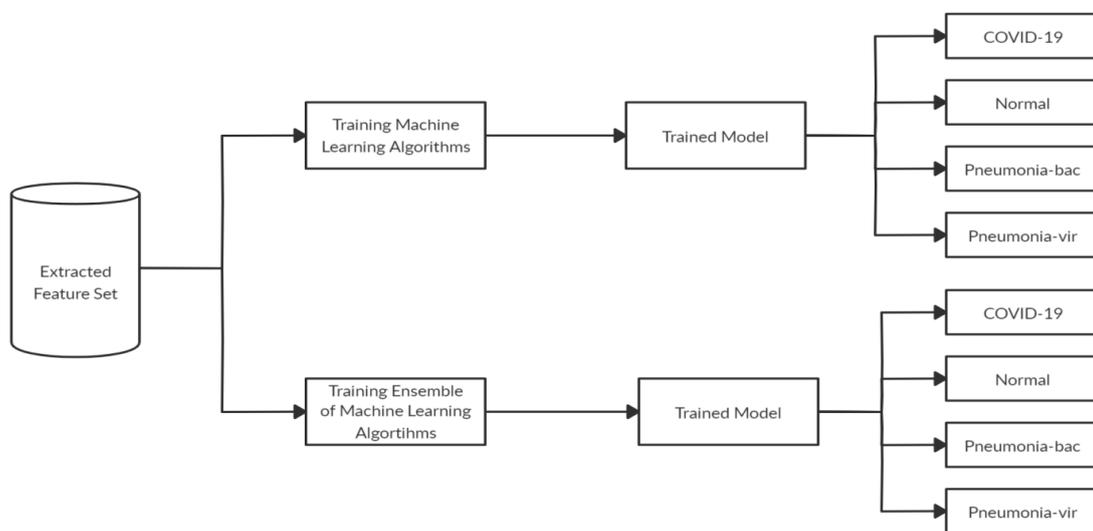


Figure 2 Phase 2 - Training of machine learning models with extracted GLCM feature sets for proposed COVID-19 detection.

Gray Level Co-occurrence Matrix (GLCM) [11]:

The GLCM is a statistical approach that gives a spatial connection of image pixels. The image textural features are calculated using the GLCM matrix, on which formulas can be applied to compute various textural features. The GLCM is a matrix that calculates frequencies of associations of values for pixels within a neighborhood, which are then normalized to probabilities.

In this paper, the GLCM is extracted first from the image. Then, the GLCM matrix is calculated in 4 directions, which are 135, 90, 45 and 0 °. In the following formulas, let p, q be the number of rows and columns of matrix respectively, and $M_{p,q}$ be the probability value recorded for the cell (p, q), and let the number of gray levels in the image be ‘N’. Several textural features are then extracted from this matrix; extracted textural features [12] are as shown in the following Eqs:

$$\text{Contrast : } \sum_{p,q=0}^{N-1} M_{p,q} (p - q)^2 \tag{1}$$

$$\text{Correlation: } \sum_{p,q=0}^{N-1} M_{p,q} \left[\frac{(p-\mu_p)(q-\mu_q)}{\sqrt{(\sigma_p^2)(\sigma_q^2)}} \right] \tag{2}$$

where, $\mu_p = \sum_{p,q=0}^{N-1} p(M_{p,q})$

$$\mu_q = \sum_{p,q=0}^{N-1} q(M_{p,q})$$

μ_p and μ_q are mean.

$$\sigma_p^2 = \sum_{p,q=0}^{N-1} M_{p,q} (p - \mu_p)^2$$

$$\sigma_q^2 = \sum_{p,q=0}^{N-1} M_{p,q} (q - \mu_q)^2$$

σ_p and σ_q are standard deviations.

$$\text{Dissimilarity : } \sum_{p,q=0}^{N-1} M_{p,q} |p - q| \quad (3)$$

$$\text{Homogeneity : } \sum_{p,q=0}^{N-1} \frac{M_{p,q}}{1+(p-q)^2} \quad (4)$$

$$\text{Angular Second Moment(ASM): } \sum_{p,q=0}^{N-1} M_{p,q}^2 \quad (5)$$

$$\text{Energy : } \sqrt{ASM} \quad (6)$$

Models of machine learning used in proposed COVID-19 detection

We used five machine learning models in the proposed technique of COVID-19 detection, which were Random Forest, Support Vector Machine (SVM), Multilayer Perceptron (MLP), Hoeffding Tree and Random Tree.

Random Forest consists of many individual decision trees operating together as a group. In the Random Forest, every individual tree provides a prediction of a class, and the class receiving the maximum number of votes becomes the prediction class of the model.

Hoeffding Tree is a type of streaming classification. It uses Hoeffding bound in the tree creation. The basic idea here is that the best attribute Hoeffding bound gives a certain level of confidence, which splits the tree.

The Support Vector Machine (SVM) model attempts to find a hyperplane in a multi-dimensional feature space which classifies all of the data points. It separates the data points of the 2 classes so that they can be classified accordingly.

A Random Tree is a tree that is figured by any random process. In this algorithm, data is constantly split based on certain parameters, which further helps in prediction.

Multilayer Perceptron (MLP) is a type of feed-forward neural network. It uses a supervised technique called backpropagation for training.

Ensembling method

The ensemble method combines several base machine learning models to produce several variations of models that can be trained on the dataset separately. This method may give better performance as compared to a single model. For example, we know that Random Forest and Random Tree are 2 individual models, while Random Forest-Random Tree means that, first, Random Forest is applied, then Random Tree is applied to the same model; this is known as an ensemble.

In this paper, we used majority voting, which means that the maximum of votes based on all the algorithms in the ensemble is considered; the class which obtains the highest number of votes is chosen.

Experimentation Environment

The experiments were done using the Python Programming Language, and classification was done using the Waikato Environment for Knowledge Analysis (WEKA) tool. Images were collected from the GitHub repository of the COVID chest X-ray image dataset [13,14]. This consists of 60 COVID-19, 70 Normal, 70 Pneumonia-bac, and 70 Pneumonia-vir chest X-ray images. The images are in JPEG format. Image samples can be seen in **Figure 3**, below.

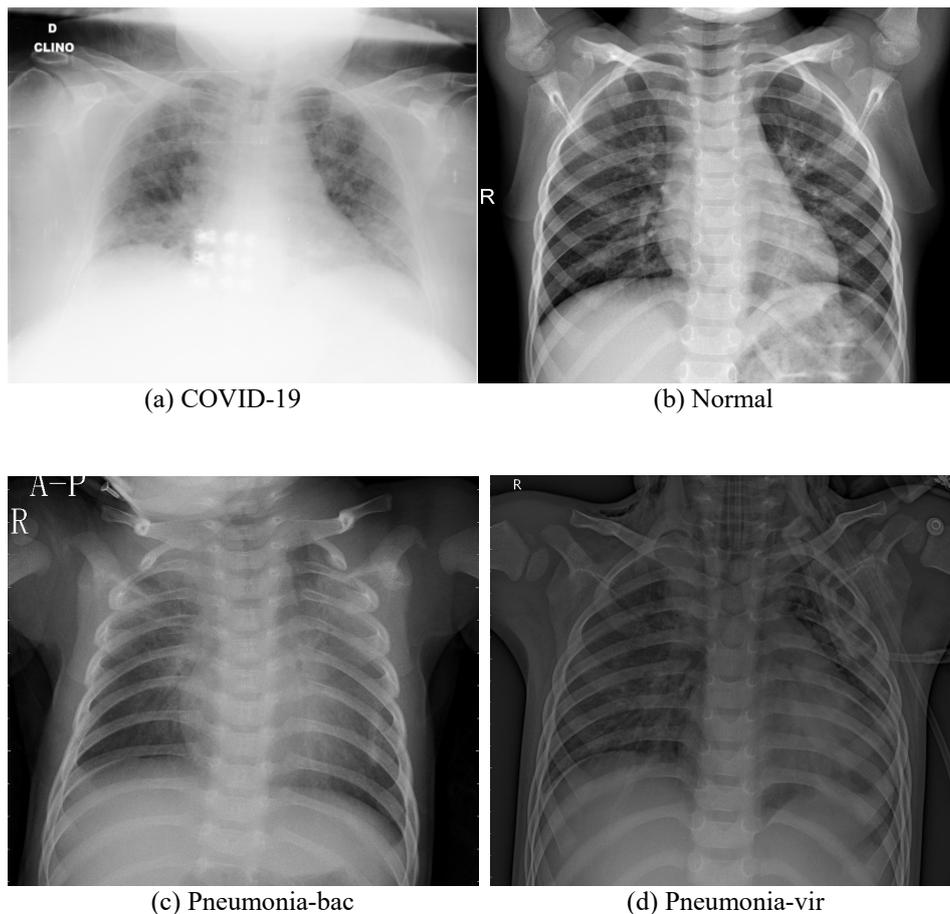


Figure 3 Image samples from the dataset.

The performance metrics which were used to assess the performance of the proposed methods were Accuracy (AC), Recall or Sensitivity (SC), Specificity (SP), Positive Predictive Value (PPV), F-measure (FM), and Mathews Correlation Coefficient(MCC), which are defined in the following Eqs. 7, 8, 9, 10, 11 and 12, respectively.

Let TPO, TNE, FPO and FNE be the count of true positive, true negative, false positive, and false negative, respectively. Then, the Eqs. are shown as follows:

$$\text{Accuracy (AC)} = \frac{TPO+TNE}{TPO + FPO + TNE + FNE} \tag{7}$$

$$\text{Recall or Sensitivity (SC)} = \frac{TPO}{TPO + FNE} \tag{8}$$

$$\text{Specificity (SP)} = \frac{TNE}{FPO + TNE} \tag{9}$$

$$\text{Positive Predictive Value (PPV)} = \frac{TPO}{TPO+FPO} \tag{10}$$

$$\text{F-Measure (FM)} = 2 \cdot \frac{SC \cdot PPV}{SC + PPV} \tag{11}$$

$$\text{Matthews Correlation Coefficient(MCC)} = \frac{TPO * TNE - FPO * FNE}{\sqrt{(TPO + FPO)(TPO + FNE)(TNE + FPO)(TNE + FNE)}} \tag{12}$$

Results and discussion

The experiment was conducted using chest X-ray images to classify them as COVID-19 or coronavirus, Normal, Pneumonia- bac (bacterial pneumonia), or Pneumonia-vir (viral pneumonia), with GLCM being used for textural feature extraction. Here, 5 machine learning models, and ensembles of those using majority voting as ensemble criteria, were used to classify the X-ray images of the chest. The observed results of the conducted experiments are discussed in this section.

Performance comparison of experimented machine learning classification models and the ensemble of machine learning models is given in **Figure 4** with the help of percentage average accuracy of classifying images of dataset. It is observed, from **Figure 4**, that better accuracy is given for the class Normal, as compared to the other classes. It can also be observed that the ensemble of Random Forest-MLP gives comparably better accuracy for almost all the classes.

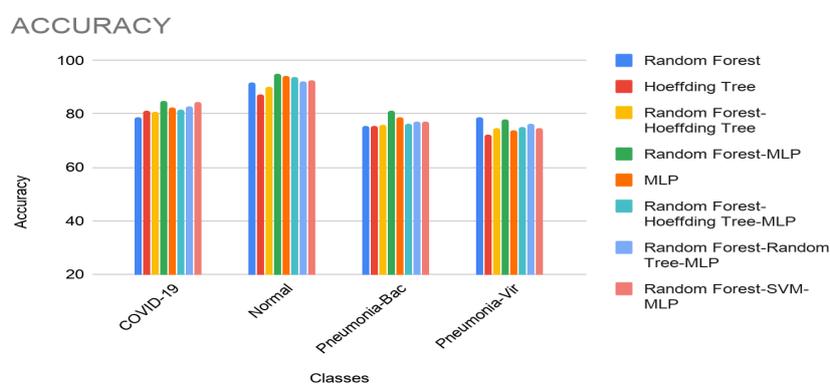


Figure 4 Comparison of machine learning classification models and ensembles of machine learning models using GLCM as feature extraction for 4 different classes with the help of percentage average accuracy.

Performance comparison of experimented machine learning classification models and ensembles of machine learning models is given in **Figure 5** with the help of precision or PPV (Positive Predictive Value) of classifying images of the dataset. It is observed, from **Figure 5**, that the ensemble of Random Forest-MLP gave better results for all the classes except for Pneumonia-vir, where Random Forest gave better results.

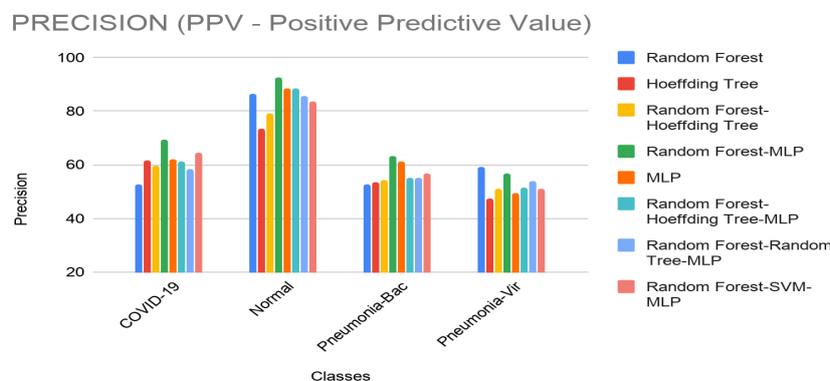


Figure 5 Comparison of machine learning classification models and ensembles of machine learning models using GLCM as feature extraction for 4 different classes with the help of precision (PPV - Positive Predictive Value).

Performance comparison of experimented machine learning classification models and ensembles of machine learning models is given in **Figure 6** with the help of recall (sensitivity) of classifying images of the dataset. It is observed, from **Figure 6**, that for classes of COVID-19, Normal, and Pneumonia-bac, the ensemble of Random Forest-MLP gave better results, while for Pneumonia-vir, the ensemble of Random Forest-SVM-MLP gave better results. Higher sensitivity is observed for the ensemble Random Forest-MLP as compared to the sensitivity values observed in other considered machine learning algorithms, as well as other ensemble combinations. The sensitivity value observed in an ensemble is better over the individual machine learning techniques, as the average prediction performance of the ensemble is improved over any contributing individual classifier due to the reduction in the variance component of prediction errors made by the individual machine learning classifiers.

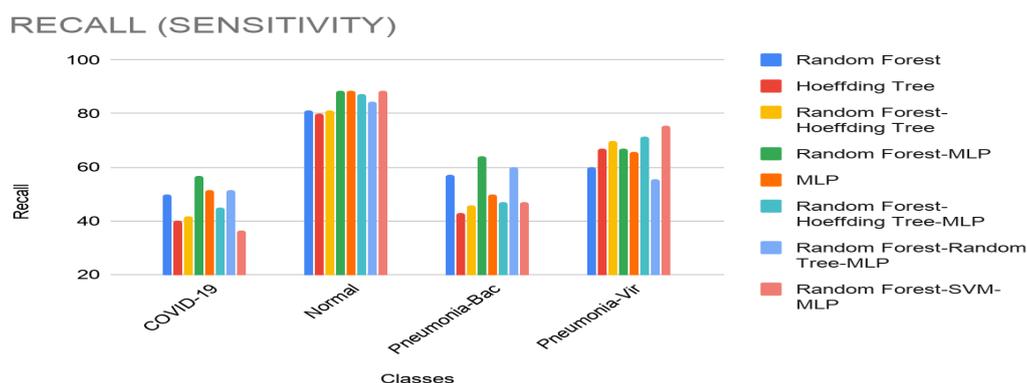


Figure 6 Comparison of machine learning classification models and ensembles of machine learning models using GLCM as feature extraction for 4 different classes with the help of percentage recall.

Performance comparison of experimented machine learning classification models and ensembles of machine learning models is given in **Figure 7** with the help of the specificity of classifying images of the dataset. It is observed, from **Figure 7**, that for classes of COVID-19, Normal, and Pneumonia-bac, the ensemble of Random Forest-MLP gave better results, while for Pneumonia-vir, the ensemble of Random Forest-SVM-MLP gave better results. Higher specificity is observed for the ensemble Random Forest-MLP as compared to the specificity values observed in other considered machine learning algorithms, as well as other ensemble combinations. The ensemble tends to minimize the variance component of estimation errors made by the individual machine learning classifiers; hence, an ensemble has shown better specificity.

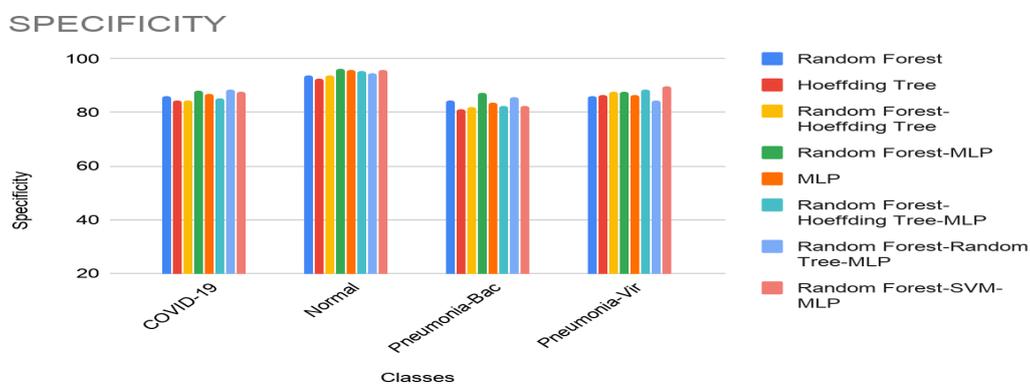


Figure 7 Comparison of machine learning classification models and ensembles of machine learning models using GLCM as feature extraction for 4 different classes with the help of percentage specificity.

Performance comparison of experimented machine learning classification models and ensembles of machine learning models is given in **Figure 8** with the help of the F-measure of classifying images of the dataset. It is observed, from **Figure 8**, that for all the classes the ensemble of Random Forest-MLP gave better results as compared to other variations. Higher F-measure values are seen in the case of the ensemble of Random Forest-MLP, indicating better suitability for COVID-19 identification. An ensemble reduces the spread or dispersion of the predictions and model performance and, hence, an ensemble gives better F-measure over the individual machine learning classifiers.

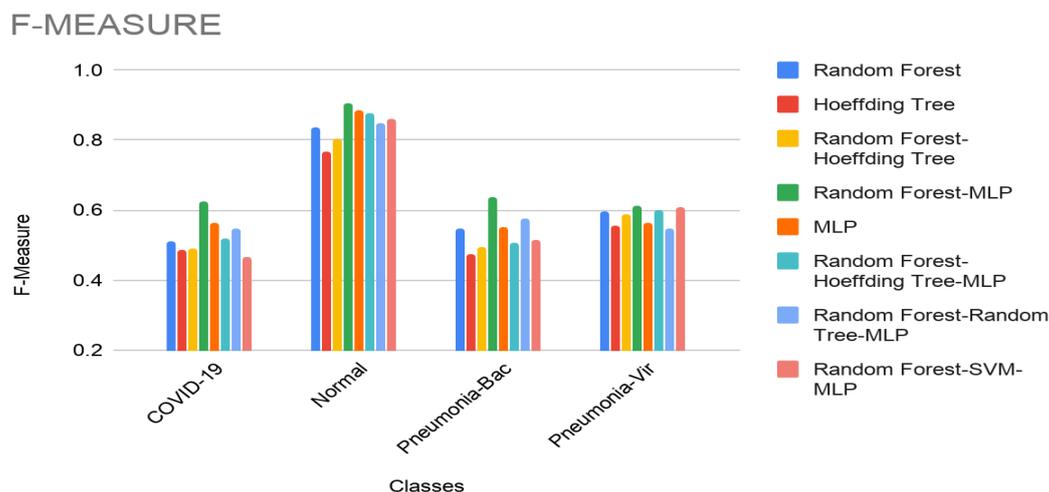


Figure 8 Comparison of machine learning classification models and ensembles of machine learning models using GLCM as feature extraction for 4 different classes with the help of F-measure.

Performance comparison of experimented machine learning classification models and ensembles of machine learning models is given in **Figure 9** with the help of the Matthews Correlation Coefficient(MCC) of classifying images of the dataset. It is observed, from **Figure 9**, that the ensemble of Random Forest-MLP gave better results comparably for all the 4 classes.

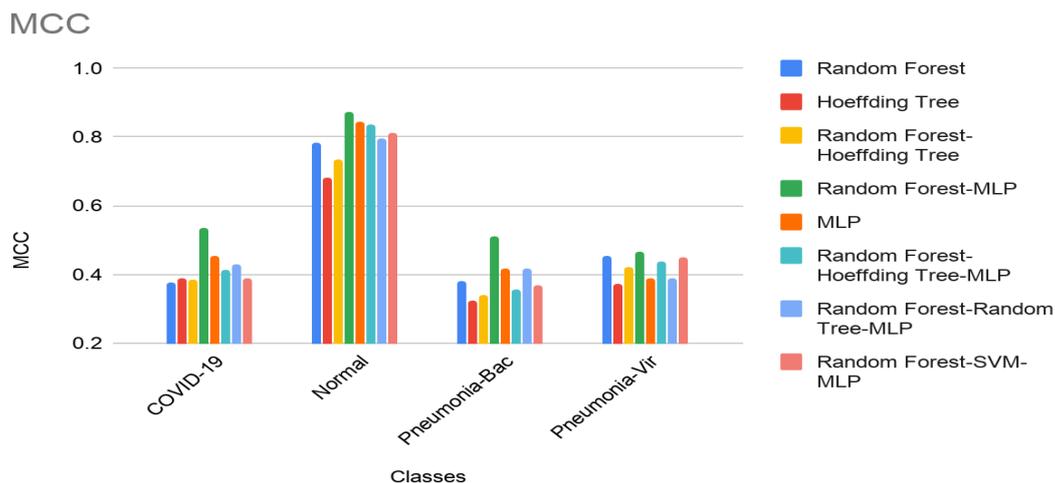


Figure 9 Comparison of machine learning classification models and ensembles of machine learning models using GLCM as feature extraction for 4 different classes with the help of MCC.

Since we were working for COVID-19 prediction, the variation of various algorithms consisted of the five machine learning algorithms used, which were Random Forest, Random Tree, Hoeffding Tree, Support Vector Machine (SVM), and Multilayer Perceptron (MLP), and also ensembles of these algorithms; from these, the top-performing variations for the prediction of COVID-19 from chest X-ray images can be seen in **Table 1**, below. The top-performing variations were chosen from all the variations by taking an average of sensitivity, specificity, PPV (Positive Predictive Value) and accuracy.

Table 1 Percentage values of Accuracy, Sensitivity, Specificity, and Positive Predictive Value (PPV) considered together in top-performing variations of the proposed COVID-19 prediction method.

Algorithm	Sensitivity	Specificity	PPV	Accuracy	All Metrics Average
Random Forest - MLP	56.7	88.235	69.387	84.814	74.784
MLP	51.7	86.818	62	82.222	70.685
Random Forest - Random Tree - MLP	51.7	88.679	58.49	82.641	70.377
Random Forest - SVM - MLP	36.7	87.61	64.705	84.615	68.407
Random Forest - Hoeffding Tree - MLP	45	85.398	61.363	81.481	68.310

A total of 270 patient X-ray images (60 COVID-19, 70 Normal, 70 Pneumonia-bac, and 70 Pneumonia-vir chest X-ray images) were considered for all experimental validations of the variations of the proposed COVID-19 identification method using GLCM features. The extracted GLCM features were used for training various machine learning classifiers and their considered majority voting ensembles. The observations made during the testing phase of experimentation for respective classifiers and ensembles were compared together for noting the machine learning ensembles to be used in top-performing variants of the proposed method. **Table 1** gives such top-performing observations (as percentage values) regarding the performance metrics used- sensitivity, specificity, PPV, and accuracy. In **Table 2**, only consideration of F-measure as a performance metric is done for noting top-performing variants of the proposed method. In **Table 3**, only consideration of MCC as the performance metric is done for noting top-performing variants of the proposed method.

Another performance metric considered was F-measure. The harmonic mean of precision and recall (sensitivity) is also known as F-measure. F-measure balances the use of precision and recall. According to our experiment, the top-performing variation according to F-measure are shown in the below **Table 2**.

Table 2 F-measure in top-performing variations of the proposed COVID-19 prediction method.

Algorithm	F-measure
Random Forest - MLP	0.624
MLP	0.564
Random Forest - Random Tree - MLP	0.549
Random Forest - Hoeffding Tree - MLP	0.519
Random Forest	0.513

Another performance metric considered was the Matthews Correlation Coefficient (MCC). This calculates the correlation coefficient between 2 classes. The higher the correlation between these classes, the better the prediction. So, according to our experiment, in the below **Table 3**, top-performing variations according to MCC are shown.

Table 3 Matthews Correlation Coefficient (MCC) in top-performing variations of the proposed COVID-19 prediction method.

Algorithm	MCC
Random Forest - MLP	0.534
MLP	0.456
Random Forest - Random Tree - MLP	0.431
Random Forest - Hoeffding Tree - MLP	0.415
Hoeffding Tree	0.389

Proposed COVID-19 identification using machine learning models with GLCM features of chest X-ray images was explored with the available number of images of chest X-ray of patients, and it was observed that the technique showed an accuracy of up to 84 % for an ensemble of machine learning models.

The main disadvantage of the technique is limited validation due to the non-availability of a sufficiently large dataset of COVID-19 patient X-ray images. Further, these preliminary results can be used for validation on a large number of images having COVID-19 patient chest X-rays.

Conclusions

This paper proposes a mechanism for the detection of COVID-19 using X-ray images of the chest. Extraction of various features was done using Gray Level Co-occurrence Matrix (GLCM); here, we considered 6 features, which were correlation, energy, dissimilarity, Angular Second Moment (ASM), contrast, and homogeneity. After extracting these features, various machine learning models, and the ensembles of machine learning models with majority voting as their ensemble types, were trained for COVID-19 identification. The performance was measured using recall (sensitivity), specificity, Matthews Correlation Coefficient (MCC), F-measure, precision (PPV - Positive Predictive Value) and accuracy. From these experimental results, we found that the ensemble of Random Forest-MLP showed the best COVID-19 identification if we took the average of sensitivity, accuracy, and precision; also, the value of F-measure of Random Forest-MLP ensemble for COVID-19 class was 0.624, with the value of MCC for COVID-19 prediction in the ensemble of Random Forest-MLP being 0.534, and was therefore comparably better for that class.

References

- [1] Covid-19 information, Available at: <https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en>, accessed June 2020.
- [2] K Roosa, Y Lee, R Luo, A Kirpich, R Rothenberg, JM Hyman, P Yan and G Chowell. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, infectious disease modelling. *Infect. Dis. Model.* 2020; **5**, 256-63.
- [3] ZL Zazhi. Epidemiology working group for NCIP epidemic response, Chinese center for disease control and prevention. *Chin. J. Epidemiol.* 2020; **41**, 145-51.
- [4] DC Cho. In the era of COVID-19: The role of spine surgeons at the Epicenter of the outbreak. *Neurospine* 2020, **17**, 337-8.
- [5] Symptoms of COVID-19, Available at: <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>, accessed June 2020.
- [6] T Ai, Z Yang, H Hou, C Zhan, C Chen, LV Wenzhi, Q Tao, Z Sun and L Xia. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology* 2020; **296**, E32-E40.
- [7] Gu Xianghong, L Pan, H Liang and R Yang. Classification of bacterial and viral childhood pneumonia using deep learning in chest radiography. *In: Proceedings of the 3rd International Conference on Multimedia and Image Processing*. Association for Computing Machinery, New York, USA. 2018, p. 88-93.

-
- [8] JTC Ming, NM Noor, OM Rijal, RM Kassim and A Yunus. Lung disease classification using GLCM and deep features from different deep learning architectures with principal component analysis. *Int. J. Integr. Eng.* 2018; **10**, 76-89.
- [9] A Narin, C Kaya and Z Pamuk. Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal. Appl.* 2021; **24**, 1207-20.
- [10] A Mangal, S Kalia, H Rajgopal, K Rangarajan, V Namboodiri, S Banerjee and C Arora. CovidAID: COVID-19 detection using chest X-ray, Available at: <https://arxiv.org/abs/2004.09803>, accessed May 2020.
- [11] M Hall-Beyer. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *Int. J. Rem. Sens.* 2017; **38**, 1312-38.
- [12] M Hall-Beyer. *GLCM texture: A tutorial v 3.0 March 2017*. University of Calgary, Canada, 2017.
- [13] JP Cohen, P Morrison, L Dao, K Roth, TQ Duong and M Ghassemi. COVID-19 image data collection: Prospective predictions are the future. *J. Mach. Learn. Biomed. Imag.* 2020; **2**, 1-38.
- [14] JP Cohen, P Morrison and L Dao. COVID-19 image data collection. Available at: <https://github.com/ieee8023/covid-chestxray-dataset>, accessed May 2020.