

Employing Texture Features of Chest X-Ray Images and Machine Learning in COVID-19 Detection and Classification

Hiam Alquran¹, Mohammad Alsleti², Roaa Alsharif³, Isam Abu Qasmieh¹,
Ali Mohammad Alqudah^{1,✉}, Nor Hazlyna binti Harun⁴

¹Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, 21163, Jordan

²The Institute of Biomedical Technology, King Hussein Medical Center, Royal Jordanian Medical Service, Amman, 11855, Jordan

³College of applied medical Sciences, Radiological Science Program, King Saud University, Jeddah, 21435, Saudi Arabia

⁴Data Science Research Lab, School of Computing, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia
ali_qudah@hotmail.com, ali_qudah@yu.edu.jo✉

Abstract

The novel coronavirus (nCoV-19) was first detected in December 2019. It had spread worldwide and was declared coronavirus disease (COVID-19) pandemic by March 2020. Patients presented with a wide range of symptoms affecting multiple organ systems predominantly the lungs. Severe cases required intensive care unit (ICU) admissions while there were asymptomatic cases as well. Although early detection of the COVID-19 virus by Real-time reverse transcription-polymerase chain reaction (RT-PCR) is effective, it is not efficient; as there can be false negatives, it is time consuming and expensive. To increase the accuracy of in-vivo detection, radiological image-based methods like a simple chest X-ray (CXR) can be utilized. This reduces the false negatives as compared to solely using the RT-PCR technique. This paper employs various image processing techniques besides extracted texture features from the radiological images and feeds them to different artificial intelligence (AI) scenarios to distinguish between normal, pneumonia, and COVID-19 cases. The best scenario is then adopted to build an automated system that can segment the chest region from the acquired image, enhance the segmented region then extract the texture features, and finally, classify it into one of the three classes. The best overall accuracy achieved is 93.1% by exploiting Ensemble classifier. Utilizing radiological data to conform to a machine learning format reduces the detection time and increase the chances of survival.

Keywords: COVID-19, COVID-19 pandemic, Respiratory infection detection, Pneumonia, K-Nearest Neighbor, Support Vector Machine, Random Forest.

Received: 24 May 2021

Accepted: 09 June 2021

Published: 21 June 2021

1 Introduction

A sudden surge in the number of patients with peculiar symptoms lead to the discovery of a novel corona virus (nCoV-2019) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). By March 2020, this new virus had spread worldwide and was declared coronavirus disease (COVID-19) pandemic by the World Health Organization (WHO). Various imaging modalities such as X-ray and Computed Tomography (CT) have been utilized in the early detection of COVID-19 related respiratory infections. Consequently, image processing techniques played a role in facilitating interpretation for diagnosis and detection of the virus related infections [13, 16, 26]. The most effective approach to monitor the spread of coronavirus is to quarantine and provide adequate care to suspected cases. Laboratory tests are used as a diagnostic device for suspected cases, but the process is time-consuming with the possibility of false positives and low supply of tools. In the early stage detection of COVID-19, some

patients may have positive pulmonary imaging manifestations, but may have negative test results in swabs of RT-PCR. Such cases are not diagnosed as suspected or confirmed [12, 15, 25]. Although the virus causes milder symptoms in about 82% of cases, the others are serious or critical [22, 28].

Successful screening and prompt medical response for the infected patients is a dire necessity in order to prevent the spread of the disease. The RT-PCR test is the gold standard screening tool used to test suspected cases [9, 23]. Although this technique is widely used, but the procedure is manual, complex, laborious and time-consuming with a positive result of just 63% [24]. Throughout the world inaccurate detection or improper count of COVID-19 patients has caused a lot of suffering, not only due to lesser tests being conducted, but also to delays in the test results [18]. Common characteristics of COVID-19 lung infection as seen radiographic images including early-stage bilateral, multi-focal, ground-glass opacities (GGO) with a

peripheral or posterior distribution, primarily in the lower lobes, and late-stage pulmonary consolidation [9, 12, 22, 23, 25, 28]. Although typical CXR helps screen suspected cases early, the images of various viral types of pneumonia are similar and findings overlap with other differentials of pneumonia and can lead to an erroneous diagnosis in the present scenario, where hospitals are crowded and run round the clock. Consequently, an incorrect diagnosis may result in non-COVID-19 viral pneumonia being falsely labeled as highly suspicious and thus delaying care and efforts for positive COVID-19 patients.

Recently, many works have been provided and here we will be surveyed the most relevant and recent works. In [15], the authors proposed a method that can reduce the requirements of manual labeling of chest CT images into COVID-19 and non-COVID-19 cases. The team plans to further develop their work on a larger scale of clinical studies. While in [12], the authors developed a classification system that is based on artificial intelligence (AI) for labeling CT images and to differentiate between COVID-19 and non-COVID-19 patients. In [25], the authors reports that they employed deep learning techniques in diagnosis lung CT and to detect COVID-19. This type of system can extract radiographic features of novel pneumonia, especially the GGO from radiographs. Also, authors in [28], built an architecture that consists of three main stages: Firstly, a random selection of ROIs. Secondly, features extraction by training the convolutional network (CNN) model, and finally, training the classifier to build a model for prediction purposes. The authors in [22], used three different neural network (NN) models i.e., ResNet50, InceptionV3, and Inception-ResNetV2 for the detection of COVID-19 lung disease. Finally, the authors in [23], proposed an NN that is based on Xception and ResNet50V2 networks to detect new cases of patients suffering from COVID-19. This network in particular achieved the best accuracy by employing various extracted features. More recent publications [5, 9, 24, 18] focus on using deep learning techniques in the detection and classification of COVID-19 disease, without any enhancement performed on the original images, whether using CT or CXR images. Also, the features that have been used in all recent studies are extracted only from deep learning techniques, besides there is no recent study that discriminated against pneumonia that results from bacteria or viruses, particularly the SARS-CoV-2 virus. This study aims to enhance the appearance of the original radiological image as well as to lay down the foundation for machine learning. The automated system will detect disease based on the features in the enhanced image that will classify chest X-ray images into normal, pneumonia, or COVID-19.

2 Materials & Methods

This section will discuss in detail the proposed methodology. The proposed methodology that has been fol-

Table 1: Distribution of X-ray Data set.

Disease Type	Number of Images
Normal	1219
Pneumonia	648
COVID-19	612

lowed in this paper is illustrated in Fig. 1.

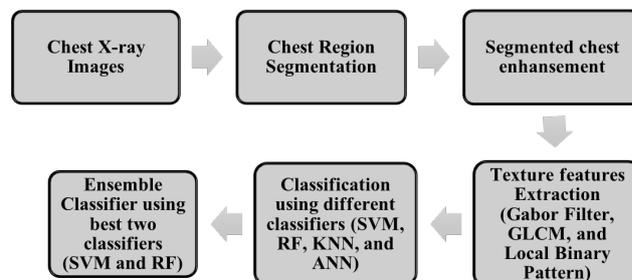


Figure 1: Effects of selecting different switching under dynamic condition.

2.1 Data Collection

The data set was collected from two websites [10, 21] with three classes: COVID-19, pneumonia, and normal. Anteroposterior (AP) view of Chest X-ray was selected from git hub website [17] with 612 COVID-19 CXR images and from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children’s Medical Center, Guangzhou with 648 Pneumonia and 1219 normal X-ray chest images, as well. All chest X-ray imaging was performed as part of patients’ routine clinical care. The total number of images is 2479. The distribution of these images is described in Table 1.

2.2 Image Dataset Preprocessing

During pre-processing of the CXR, initially, the image is converted into binary by auto-thresholding. Then the largest rectangle detection is applied to extract the region which contains the lungs region [1]. Fig. 2 (a) shows the original chest X-ray image and (b) its corresponding extracted region. Then the chest region segmented image is enhanced using gray-level transformation enhancement methods. It starts by using a 3×3 median filter. Followed by the image sharpening using the unsharp masking method. The intensity values are adjusted to increase the contrast of the pixels in the image of the lung field. If r is the original gray value and $T(r)$ is the new gray value. σ is the standard deviation and has a value of 240. Then in order to improve the appearance of the enhanced image transformation function, which has Gaussian distribution behavior, is applied:

$$T(r) = 255 \times e^{-\frac{(128-r)^2}{2\sigma^2}} \quad (1)$$

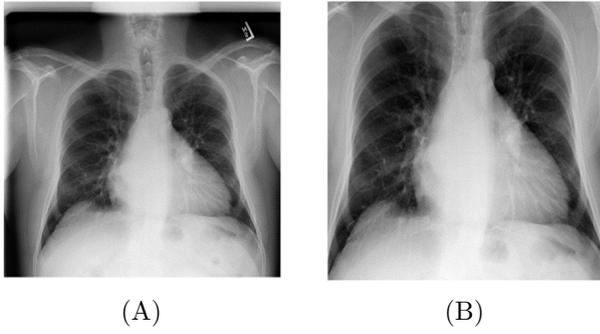


Figure 2: (a) Original chest X-ray Image (b) The corresponding largest lungs region.

Fig. 3 shows the relation between the original and the new intensity values where (a) shows the segmented image and (b) depicts the resultant image after intensity transformation and color-mapping to pink.

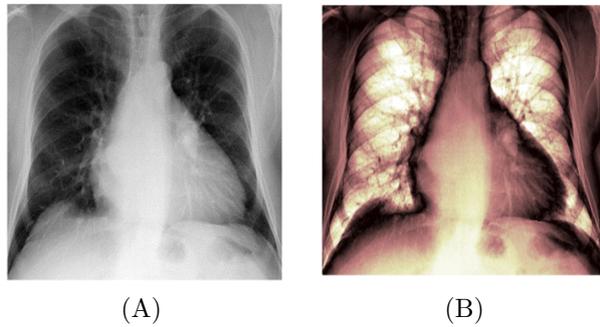


Figure 3: (a) Segmented image (b) Resultant image after intensity transformation.

Fig. 4 represents the transformation function that is applied to the segmented image. The automated segmented image is compared with that which is segmented manually; the segmentation accuracy is 98.3%.

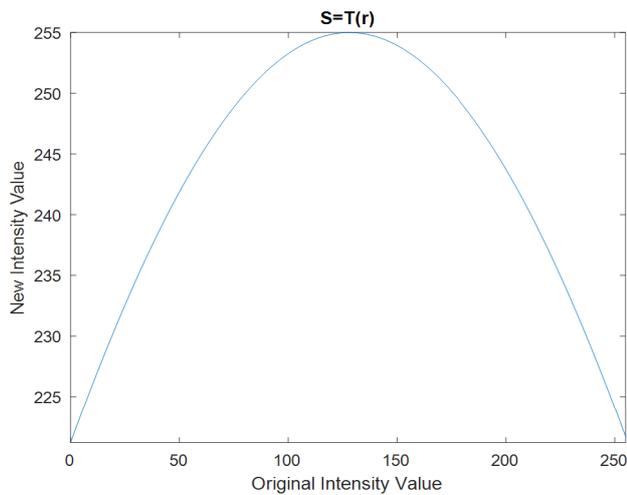


Figure 4: Transformation function.

2.3 Texture Features Extraction

Texture is commonly used in many computer vision applications. It can be defined as a measure of coarseness, contrast, directionality, line-likeness, regularity, and roughness. It appears as a grouping of similarities or repetitive arrangement of pixels. Therefore, texture features play the main role in characterizing the nature of human tissue structure. Engineers and researchers have proposed many methods for texture feature extraction. In this study, statistical approaches such as gray level co-occurrence matrix and local binary pattern were employed to transform approaches such as the Gabor decomposition method in extracting the features of interest [14].

2.3.1 Local Binary Pattern

Local binary pattern (LBP) operator measures local contrast of the image. Lahdenoja [27] employed LBP operator besides structural texture analysis and statistical features of the target tissue. LBP describes texture with the smallest primitives called textons [2]. For each pixel in the image, thresholding creates a binary code that is based on comparing the central pixel value with its eight closest neighbors. If the neighbor pixel value is greater than the central pixel value, it is replaced by one, otherwise, it is changed to zero. The same procedure is followed for all cells. After that, a histogram is calculated and normalized. The result is a vector of 59 features, where the standard deviation and mean are computed from the LBP vector [30]. LBP is one of the most applicable approaches in many applications due to its simplicity and performance.

2.3.2 Gabor Filter

Gabor filter is used in analyzing the texture of images in multiresolution representations. Many studies have demonstrated the power of the Gabor filter in detecting texture frequency and orientations. The 2D Gabor filter is described at spatial index (x, y) as [29].

$$g_{\lambda, \theta, \sigma, \psi}(x, y) = e^{-\frac{x_{\theta}^2 + \gamma^2 y_{\theta}^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda} x_{\theta} + \psi\right) \quad (2)$$

$$x_{\theta} = x \cos(\theta) + y \sin(\theta) \quad (3)$$

$$y_{\theta} = -x \sin(\theta) + y \cos(\theta) \quad (4)$$

$$\lambda = \frac{1}{f} \quad (5)$$

where σ is the standard deviation of the Gaussian function in the x and y directions, and θ indicates the orientation of the filter. Mean and standard deviation of the magnitude and phase are exploited as features for classification purposes.

2.3.3 Gray Co-Occurrence Matrix (GLCM)

GLCM describes the frequency of one gray level that appears in a specified spatial linear relationship with another gray level within the region under investigation. The texture feature utilizes the contents of the GLCM to quantify the variation in specific intensity of pixels. Traditionally, computation of this matrix is based on two parameters: relative distance between the pixel paired and their relative orientation. It is calculated in four directions such as; 0° , 45° , 90° , and 135° . Various features like autocorrelation, cluster prominence, cluster shade, contrast, correlation, difference entropy, difference variance, dissimilarity, energy, entropy, homogeneity, inverse difference, sum average, sum entropy, sum of squares, sum variance and inverse difference moment normalized have been extracted from GLCM [20].

2.4 Principal Component Analysis (PCA)

PCA is a well-known algorithm that has been used in dimension reduction apart from its ability to extract the most significant features [6]. It was applied to the resultant GLCM features. Furthermore, the 8 most independent features are Cluster Shade, the Sum of Squares, Correlation Sum Average, Entropy, Sum of Entropy, and Inverse Difference Normalized Moment. Therefore, 14 features that have been fed to the classifiers are 8 most independent GLCM features, Mean of Gabor Magnitude, Standard deviation of Gabor Magnitude, Mean of Gabor phase, Standard deviation of Gabor phase, Mean of LPB features, and Standard deviation of LBP features.

2.5 Classification

The main part of a statistical pattern recognition system is the classifier type. There are quite a few classification algorithms in machine learning, such as Naïve Bayes, C4.5 tree, K-NN, K-means, Neural Networks (NN), SVM, AdaBoost and Random Forest (RF). The classifiers used in this study are SVM, RF, K-NN, and Artificial Neural Network (ANN). MATLAB® 2019b has been used to design SVM, RF, NN, and K-nearest neighbor classifiers. The classification begins by distinguishing between Pneumonia, COVID-19, and Normal chest X-ray images for each type of previously mentioned classifiers. Then, the classification is carried out to a combination of two classes (such as Normal with Pneumonia, Normal with COVID-19, and Pneumonia with COVID-19) for each classifier individually. The results are compared to conclude the best classifier that gives high accuracy for each specific class. Lastly, an Ensemble classifier utilizes the benefits of the recorded results to build as series classifiers to get the highest discrimination for each class. This classifier is built from two stages, the first stage to separate normal from abnormal images, and the second stage classifies the abnormal cases into Pneumonia or COVID-19.

2.5.1 Support Vector Machine (SVM)

Support Vector Machine classifier is a binary supervised classifier, and it can be extended to multiclass classification applications. It is optimized by finding hyperplanes that maximize the separation region between the margins of two classes. This can be obtained by mapping the used features into a higher-dimensional space using kernel functions such as linear, polynomial, and radial basis function (RBF). If the kernel is linear, then it is called linear SVM and it is used when the data is linearly separable. Otherwise, other kernels have been used such as Polynomial, RBF etc., to discriminate between different classes [8]. In this paper SVM classifier is designed with 3rd order polynomial kernel function.

2.5.2 Random Forest (RF)

Random forests classifier consists of multiple decision trees. Each node in the tree includes a group of training instances and a predictor. A random subset of features is selected at each attribute split based on the bagging method. The trees keep growing to reach a specified depth, and a class voting is determined after generating many trees [3].

2.5.3 k-Nearest Neighbor (KNN)

The supervised KNN classifier algorithm was proposed by Fix and Hodges in 1951. It classifies a data point based on its neighbors' class. The classification results are specified based on k value of the nearest neighbor, k value was specified as 1. To classify the new sample based on its attribute vector, the closest k samples are selected from the training data. Accordingly, the new vector is addressed to it by looking at the classes in which the candidate's samples be classified [4].

2.5.4 Artificial Neural Network (ANN)

An NN is a set of connected input/output units in which each connection has a weight associated with it. An ANN is a computational model that consists of many simple connected units (neurons) that work in parallel. The basic elements of a NN are the units, the connections between units, the weights, and the thresholds [19]. All data must be normalized and all values of attributes in the database are changed to include values in the interval $[0,1]$ or $[-1,1]$.

2.5.5 Ensemble Classifier

Ensemble classifier is based on constructing a set of classifiers and classifying the new data by considering the weighted vote of their prediction [11]. Fig. 5 shows the Ensemble classifier that has been designed for this study.

After pre-processing techniques and extracting features, those features are fed to ensemble classifier to the first stage which is RF classifier, if the output is

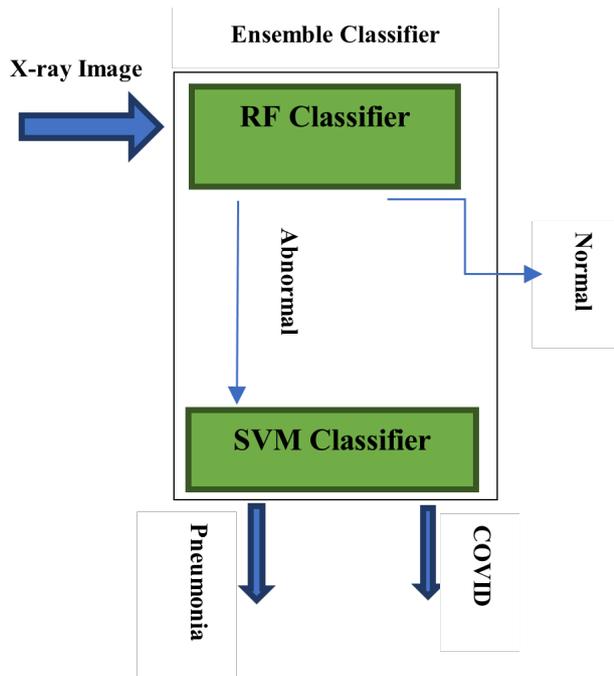


Figure 5: Ensemble Classifier architecture.

normal, then the diagnosis is normal, but if it is abnormal, the feature is subjective to second stage which is SVM classifier. This classifier discriminates if this abnormality comes from COVID-19 virus or it is from other types of viral or bacterial pneumonia.

3 Results

The performance of the selected classifier is evaluated using sensitivity and specificity statistical measures. Sensitivity describes the ability to identify the True Positive cases, while Specificity describes the True Negative cases. Accuracy represents the percentage of correct or true results. The significant value for all is 1 [7].

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + F_N + T_N + F_P} \quad (6)$$

$$\text{Sensitivity} = \frac{T_P}{T_P + F_N} \quad (7)$$

$$\text{Specificity} = \frac{T_N}{T_N + F_P} \quad (8)$$

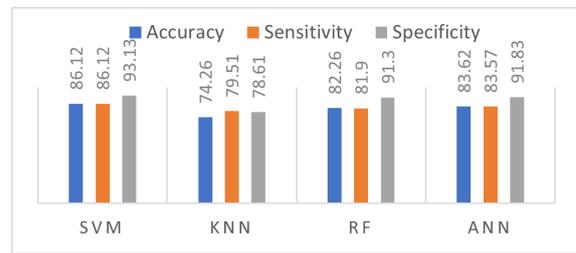
$$\text{Precision} = \frac{T_P}{T_P + F_P} \quad (9)$$

$$\text{F1-Score} = 2 \frac{\text{Sensitivity} \cdot \text{Precision}}{\text{Sensitivity} + \text{Precision}} \quad (10)$$

The recognition between the three classes is performed using various scenarios to conclude the best classifier type that gives the higher accuracy value when tested with new images. The best classifier candidate will be the user interface using the Graphical User Interface (GUI) to classify input chest X-ray image into three classes. The adopted scenarios are as follows:

		Target Class					
		Normal		Pneumonia		COVID-19	
Normal	SVM	197	SVM	26	SVM	5	
	KNN	185	KNN	32	KNN	11	
	RF	198	RF	28	RF	6	
	ANN	178	ANN	11	ANN	3	
Pneumonia	SVM	15	SVM	144	SVM	22	
	KNN	27	KNN	129	KNN	40	
	RF	11	RF	138	RF	21	
	ANN	27	ANN	146	ANN	10	
COVID-19	SVM	5	SVM	24	SVM	137	
	KNN	5	KNN	33	KNN	113	
	RF	8	RF	28	RF	137	
	ANN	12	ANN	30	ANN	151	

(a)



(b)

Figure 6: (a) Confusion matrices (b) Test Phase Performance Evaluation of Four Classifiers.

1. Three classes classification using various classifiers.
2. Two classes classification using four types of classifiers.
3. Three classes classification using Ensemble Classifier.

3.1 All Classes Classification

The first scenario is classifying the input image into three categories i.e., Normal, COVID-19, and Pneumonia. Then the data is divided into 70% training and 30% testing. Fig. 6 (a) and (b) shows the confusion matrix in test phase for all selected classifiers, the overall accuracy in the test phase, sensitivity for each class in all classifiers, and the specificity for each class in all classifiers types, respectively. SVM is the best of all four. Using these figures ANN has the highest sensitivity in recognizing Normal and COVID-19 cases, while RF has the highest recall for Pneumonia cases. ANN achieved the highest specificity in recognizing Normal and Pneumonia classes, but the best specificity for COVID-19 recognition results are obtained using RF or SVM classifiers.

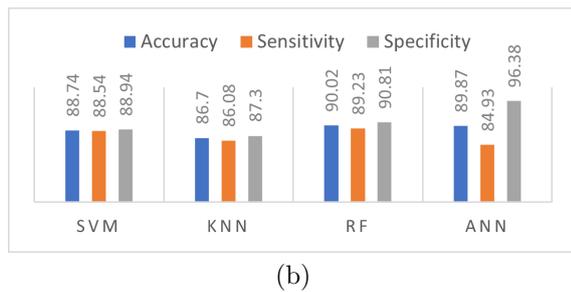
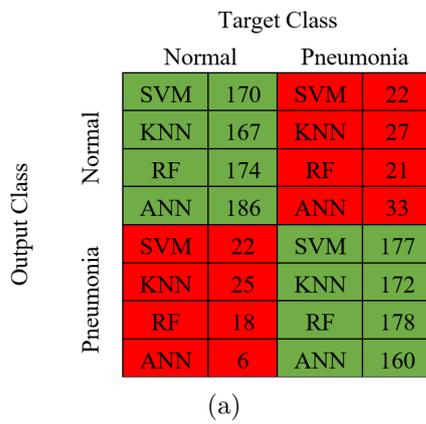


Figure 7: (a) Confusion matrices (b) Test Phase Performance Evaluation of Four Classifiers.

3.2 Two Classes Classification

The classification procedure is applied for various classifiers to discriminate between two classes.

3.2.1 Normal vs Pneumonia Classes

Fig. 7 (a) and (b) shows the confusion matrix and overall accuracy in the test phase, sensitivity for two classes in all classifiers, and the specificity for each class in all classifier types, respectively. ANN is the best of all four. Pneumonia class has higher specificity using ANN and Normal class obtained higher sensitivity using ANN class.

3.2.2 Normal vs COVID-19

Fig. 8 (a) and (b) shows the confusion matrix and overall accuracy in the test phase, sensitivity for two classes in all classifiers, and the specificity for each class in all classifiers types, respectively. RF is the best among all four classifiers. RF classifier gives higher sensitivity in Normal class and higher specificity in COVID-19 class.

3.2.3 Pneumonia vs COVID-19

Fig. 9 (a) and (b) shows the confusion matrices and overall accuracy in the test phase, sensitivity for two classes in all classifiers, and the specificity for each class in all classifiers types, respectively. RF is the best among all four classifiers.

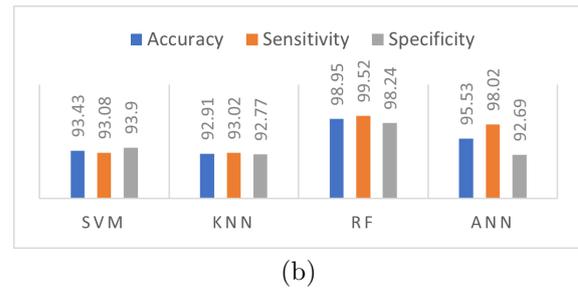
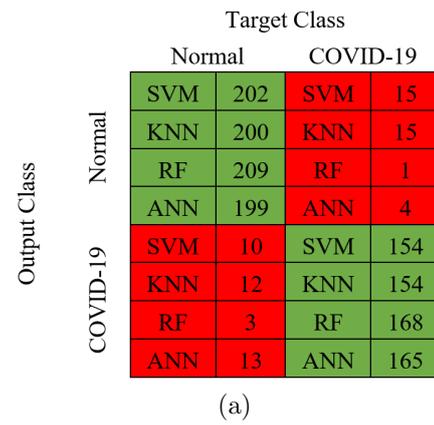


Figure 8: (a) Confusion matrices (b) Test Phase Performance Evaluation of Four Classifiers.

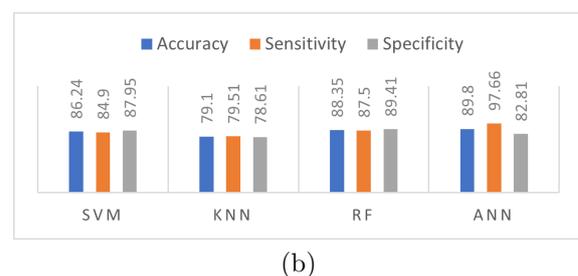
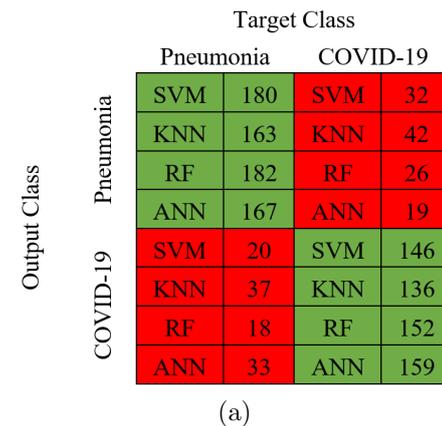


Figure 9: (a) Confusion matrices (b) Test Phase Performance Evaluation of Four Classifiers.

3.3 Ensemble Classifier for Three Classes

The last scenario is performed by using the Ensemble classifier to design according to the previous results of the two classes classification, and therefore, built in two stages. The first stage is using RF to discriminate between Normal and Abnormal cases i.e., Pneumonia and COVID-19. It shows that RF classifier has the

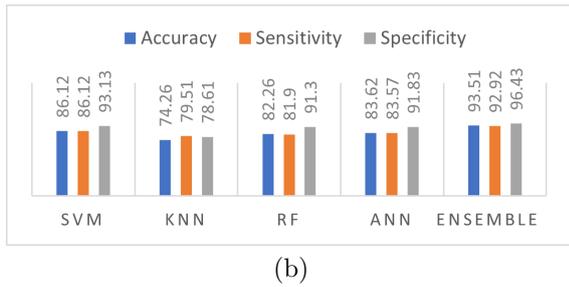
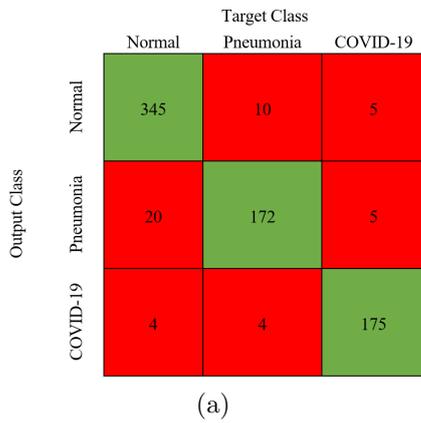


Figure 10: (a) Confusion matrices (b) Test Phase Performance Evaluation of Four Classifiers.

highest accuracy among all classifiers in discriminating Normal cases. The second stage takes the abnormal cases to classify them into Pneumonia or COVID-19. SVM classifier with order 3 is employed in the second stage because it achieved the highest sensitivity in discriminating COVID-19 cases from Pneumonia cases. This cascade classifier reaches overall accuracy of 93.1% and the test accuracy reaches 88% and training is 95.3%. The specificity for Normal cases in the test phase is 89.2%, Pneumonia is 82.4% and COVID-19 is 91.2%, the sensitivity in the test phase for normal is 91.3%, Pneumonia is 89.5% and COVID-19 is 88.1% The overall specificity for Normal is 95.4%, Pneumonia is 88.3%, and COVID-19 is 93.9%. The overall sensitivity for Normal is 96.4%, Pneumonia is 89.5% and COVID-19 is 90%. Fig. 10 (a) and (b) shows the confusion matrices and the overall accuracy in the test phase, sensitivity for two classes in all classifiers, and the specificity for each class in all classifiers types, respectively. The ensemble classifier (two stages) illustrates the best performance among all five classifiers.

The accuracy indicates the classification in all classes is precise while the sensitivity describes the positivity of the classifier. In the case of COVID-19, the sensitivity is the most important of all performance indices, because it represents the presence of the widespread Coronavirus infection hotspots in chest X-ray images. Therefore, the sensitivity of the Ensemble classifier for COVID-19 cases is the key point in concluding the best classifier for building an automated system. Fig. 11 illustrates the designed system. This type of system can be used as a standalone software and as well as a preliminary decision-making process tool for the critical

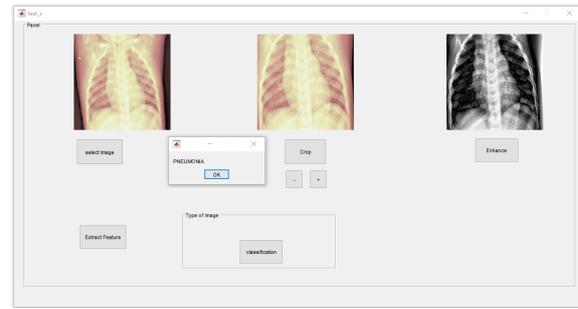


Figure 11: Graphical User Interface (GUI) of the designed system to classify the chest X-ray images.

Table 2: Results for each Scenario.

Scenario #	Classifier	Accuracy %	Sensitivity %	Specificity %	Precision %	F1-Score %
1	SVM	83.13	82.83	91.65	82.84	82.79
	KNN	74.26	73.93	87.19	73.55	73.68
	RF	82.26	81.90	91.30	81.97	81.77
	ANN	83.62	83.57	91.83	84.05	83.51
2	SVM	88.74	88.54	88.94	88.54	88.54
	KNN	86.70	86.08	87.30	86.97	86.53
	RF	90.03	89.23	90.82	90.63	89.92
	ANN	89.87	84.93	96.38	96.87	90.51
3	SVM	93.43	93.08	93.90	95.28	94.17
	KNN	92.91	93.02	92.77	94.33	93.67
	RF	98.95	99.52	98.24	98.58	99.05
	ANN	95.53	98.02	92.69	93.86	95.90
4	SVM	86.24	84.90	87.95	90.00	87.37
	KNN	79.10	79.51	78.61	81.50	80.49
	RF	88.35	87.50	89.41	91.00	89.21
	ANN	86.24	89.78	82.81	83.50	86.52
5	Ensemble	93.51	92.92	96.43	93.52	93.19

or admitted cases until the laboratory tests are completed. It helps in quick diagnosis besides other tests that must be done to make an accurate judgment about the patient.

Fig. 11 clearly shows that the input image is cropped to make the chest region more pronounced and to reduce the appearance of other parts in the image. Therefore, the resultant image appearance is enhanced and ready for extracting texture features and other features followed by the classification process using the Ensemble classifier. The process is automated with all classes overall accuracy of 93.1%. Table 2 shows a full performance evaluation of the proposed systems.

4 Discussion

This proposed method is based on the observation that the presence of COVID-19 effects on the features of pulmonary regions which is different from other types of Pneumonia. Therefore, utilization texture features apart from RT-PCR test will give us a good indication on the presence or absence coronavirus infected lung tissue. This is the first study that considers lung tis-

Table 3: Comparison with Literature.

Ref#	Dataset	Methods	Classes	Accuracy
[15]	60 3D CT lung scans	Supervised deep learning methods	COVID vs Non COVID	96.2%
[12]	157 international patients (China and U.S)	deep learning based automated CT image analysis tools	COVID vs Non COVID	99.6%
[25]	275 CT chest images	deep learning-based CT diagnosis system	COVID vs Non COVID	99%
[28]	1065 CT images	modified the Inception transfer-learning model	COVID vs Non COVID	89.5%
[22]	7065 chest X-ray images	re-trained convolutional neural network-based models	COVID vs Non COVID	96.1%
[23]	23936 x-ray images	concatenation of Xception and ResNet50V2 networks	COVID, Pneumonia, Normal	91.4%
[9]	1369 CT images	deep learning	COVID vs Non COVID	97%
[24]	381 X-ray images	Hybrid models (Deep learning and SVM)	COVID, Pneumonia, Normal	98.66%
[18]	4352 chest CT Images	deep learning model	COVID vs Non COVID	95%
[5]	71 chest X-ray images	Hybrid artificial intelligence models	COVID vs Non COVID	95.2%
This Paper	2479 chest X-ray images	Texture Features with machine learning models	COVID, Pneumonia Normal	93.1%

sue descriptors when it is compared with literature as shown in Table 3. All previous studies were focused on graphical features of chest images that are extracted automatically using deep learning techniques.

As clear from previous table the proposed approach obtained a high accuracy when it was compared with all existence methods. In order to further improve accuracy of the model and its robustness, a number of images will be added to the training data set. This model can be used as mobile application to detect COVID-19 within seconds, that may cause to reduce the workload of physicians in hospitals. This study varies significantly from all previous study by concerning on the affected region of lung tissue beside to employing various machine learning algorithms to achieve the best model that can distinguish between COVID-19 cases and other types of pneumonia cases.

5 Conclusion

The coronavirus pandemic has challenged researchers to venture for a new, precise, and faster method besides the traditional methods to investigate the presence or absence of COVID-19 affected pulmonary hotspots and to discriminate it from other types of Pneumonias. The proposed method starts from chest segmentation, then enhancement of the excluded region to make the classification task more accurate. After that, fourteen texture features are extracted from the enhanced image; mean and standard deviation of magnitude and phase of Gabor filter, mean and standard deviation of LBP vector, Sum variance, Cluster Shade, Sum of squares, Autocorrelation, Sum average, Entropy, Sum of Entropy, and Inverse difference moment normalized. Various scenarios have been developed for this study to classify the CXR image, and then the designed au-

tomated system used Ensemble classifier owing to its ability to distinguish between three classes i.e., Normal, Pneumonia, and COVID-19. Not just that but Ensemble classifier exhibits higher accuracy, compared to other tested classifiers in this study as well as sporting superior sensitivity to the presence of the coronavirus infection in the lung tissue. This study is the first study that employs texture features to test the difference between the Pneumonia that is resulted from other types of viruses or bacteria and that is infected with the SARS-CoV-2. The result of this study demonstrates that the impact of the coronavirus on lung tissue is different from other types of Pneumonia. The suggested model is accurate, fast-performing, and precise in discriminating between three chest conditions i.e., Normal, Pneumonia, and COVID-19. The future work will focus on using a huge dataset of chest CT 3D images and then employing 3D texture features to classify the chest images to obtain better results. The proposed method will be improved using 3D segmentation of the lungs region and applying more enhancement techniques to achieve accuracy.

References

- [1] ABU-QASMIEH, I., AND AL-QURAN, H. Unrestricted lr detection for biomedical applications using coarse-to-fine hierarchical approach. *IET Image Processing* 12, 9 (2018), 1639–1645.
- [2] ALQUDAH, A., AND ALQUDAH, A. M. Sliding window based support vector machine system for classification of breast cancer using histopathological microscopic images. *IETE Journal of Research* (2019), 1–9.
- [3] ALQUDAH, A. M. Ovarian cancer classification using serum proteomic profiling and wavelet fea-

- tures a comparison of machine learning and features selection algorithms. *Journal of Clinical Engineering* 44 (2019), 165–173.
- [4] ALQUDAH, A. M. Towards classifying non-segmented heart sound records using instantaneous frequency based features. *Journal of Medical Engineering & Technology* 43, 7 (2019), 418–430.
 - [5] ALQUDAH, A. M., QAZAN, S., ALQURAN, H., QASMIEH, I. A., AND ALQUDAH, A. Covid-19 detection from x-ray images using different artificial intelligence hybrid models. *Jordan Journal of Electrical Engineering* 6, 2 (2020), 168–178.
 - [6] ALQURAN, H., ABU-QASMIEH, I., KHRESAT, S., YOUNES, A., AND ALMOMANI, S. Weight estimation for anesthetic administration using singular value decomposition and template matching for supine subject of different obesity levels. *Health and Technology* 8 (2018), 265–269.
 - [7] ALQURAN, H., ET AL. Ecg classification using higher order spectral estimation and deep learning techniques. *Neural Network World* (2019), 207–219.
 - [8] ALQURAN, H., SHAHEEN, E., O’CONNOR, J. M., AND MAHD, M. Enhancement of 3D modeling and classification of microcalcifications in breast computed tomography (BCT). In *Medical Imaging 2014: Image Processing* (2014), S. Ourselin and M. A. Styner, Eds., vol. 9034, International Society for Optics and Photonics, SPIE, pp. 799 – 807.
 - [9] AMYAR, A., MODZELEWSKI, R., LI, H., AND RUAN, S. Multi-task deep learning based ct imaging analysis for covid-19 pneumonia: Classification and segmentation. *Computers in Biology and Medicine* 126 (2020), 104037 – 104037.
 - [10] COHEN, J. P., ET AL. Covid-19 image data collection: Prospective predictions are the future, 2020.
 - [11] DIETTERICH, T. G. Ensemble methods in machine learning. In *Proceedings of the First International Workshop on Multiple Classifier Systems* (Berlin, Heidelberg, 2000), MCS ’00, Springer-Verlag, p. 1–15.
 - [12] GOZES, O., ET AL. Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis, 2020.
 - [13] HAMED, A. Image processing of corona virus using interferometry. *Optics and Photonics Journal* 06 (2016), 75–86.
 - [14] HE, D.-C., WANG, L., AND GUIBERT, J. Texture feature extraction. *Pattern Recognition Letters* 6, 4 (1987), 269–273.
 - [15] HU, S., ET AL. Weakly supervised deep learning for covid-19 infection detection and classification from ct images. *IEEE Access* 8 (2020), 118869–118883.
 - [16] JIN, Y. J., ET AL. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-ncov) infected pneumonia (standard version). *Military Med Res* 7 (2020), 4.
 - [17] KERMANY, D., ZHANG, K., AND GOLDBAUM, M. Labeled optical coherence tomography (oct) and chest x-ray images for classification, mendeley data [online], 2018. doi: 10.17632/rschjbr9sj.2.
 - [18] LI, L., ET AL. Using artificial intelligence to detect covid-19 and community-acquired pneumonia based on pulmonary ct: Evaluation of the diagnostic accuracy. *Radiology* 296, 2 (2020), E65–E71.
 - [19] MAGESHKUMAR, C., ET AL. Gabor features and lda based face recognition with ann classifier. *2011 International Conference on Emerging Trends in Electrical and Computer Technology* (2011), 831–836.
 - [20] MOHANAIHAH, P., SATHYANARAYANA, P., AND GURUKUMAR, L. Image texture feature extraction using glcm approach. *International journal of scientific and research publications* 3 (2013), 1–5.
 - [21] MOONEY, P. Chest x-ray images (pneumonia) [online]. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
 - [22] NARIN, A., KAYA, C., AND PAMUK, Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Analysis and Applications* (2021).
 - [23] RAHMZADEH, M., AND ATTAR, A. A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2. *Informatics in Medicine Unlocked* 19 (2020), 100360.
 - [24] SETHY, P. K., AND BEHERA, S. K. Detection of coronavirus disease (covid-19) based on deep features. *Preprints* (2020), 2020030300.
 - [25] SONG, Y., ET AL. Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2021), 1–1.
 - [26] STOECKLIN, S. B., ET AL. First cases of coronavirus disease 2019 (covid-19) in france: surveillance, investigations and control measures, january 2020. *Eurosurveillance* 25 (2020).
 - [27] WAN, S., ET AL. Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy. *Medical Image Analysis* 38 (2017), 104–116.
 - [28] WANG, S., ET AL. A deep learning algorithm using ct images to screen for corona virus disease (covid-19). *European Radiology* (2021), 1–9.
 - [29] ZHANG, D., WONG, A., INDRAWAN, M., AND LU, G. Content based image retrieval using gabor texture features. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2000), 13–15.
 - [30] ZHAO, Y., JIA, W., HU, R.-X., AND MIN, H. Completed robust local binary pattern for texture classification. *Neurocomputing* 106 (2013), 68–76.