Identification of Pneumonia Disease Applying an Intelligent Computational Framework Based on Deep Learning and Machine Learning Techniques

Yar Muhammad, Mohammad Dahman Alshehri, Wael Mohammed Alenazy, Truong Vinh Hoang, and Ryan Alturki

1Department of Computer Science, Abdul Wali Khan University Mardan, Mardan, Khyber Pakhtunkhwa, Pakistan
2Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
3Department of Self Development Skills, CFY Deanship King Saud University, Riyadh, Saudi Arabia
4Department of Information Technology Specialization, FPT University, Hoa Lac High Tech Park, Hanoi, Vietnam
5Department of Information Science, College of Computer and Information Systems, Umm Al-Qura University, Makkah, Saudi Arabia

Correspondence should be addressed to Mohammad Dahman Alshehri; alshehri@tu.edu.sa

Received 23 March 2021; Revised 30 April 2021; Accepted 7 May 2021; Published 18 May 2021

Academic Editor: Muhammad Usman

Copyright © 2021 Yar Muhammad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pneumonia is a very common and fatal disease, which needs to be identified at the initial stages in order to prevent a patient having this disease from more damage and help him/her in saving his/her life. Various techniques are used for the diagnosis of pneumonia including chest X-ray, CT scan, blood culture, sputum culture, fluid sample, bronchoscopy, and pulse oximetry. Medical image analysis plays a vital role in the diagnosis of various diseases like MERS, COVID-19, pneumonia, etc. and is considered to be one of the auspicious research areas. To analyze chest X-ray images accurately, there is a need for an expert radiologist who possesses expertise and experience in the desired domain. According to the World Health Organization (WHO) report, about 2/3 people in the world still do not have access to the radiologist, in order to diagnose their disease. This study proposes a DL framework to diagnose pneumonia disease in an efficient and effective manner. Various Deep Convolutional Neural Network (DCNN) transfer learning techniques such as AlexNet, SqueezeNet, VGG16, VGG19, and Inception-V3 are utilized for extracting useful features from the chest X-ray images. In this study, several machine learning (ML) classifiers are utilized. The proposed system has been trained and tested on chest X-ray and CT images dataset. In order to examine the stability and effectiveness of the proposed system, different performance measures have been utilized. The proposed system is intended to be beneficial and supportive for medical doctors to accurately and efficiently diagnose pneumonia disease.

1. Introduction

Chronic diseases and epidemics have taken the lives of a large number of people and created numerous crises in the countries, which takes a long time for a country to recover from the loss caused by both of these major outbreaks. Some diseases that ascend in a specific time period within a population are termed outbreaks and epidemics [1]. Epidemic means the incidence of more cases of a disease at a particular period of time than expectations in an area, country, or group of people. The term outbreak is considered to be local and does not cause people to be panic.

Pneumonia is an infective disease that inflames the air sacs in a single or both lungs caused by fungi, bacteria, and viruses [2]. In addition, the pulmonary alveoli are affected very badly by the lung infection, the small balloon shape bags at the bottom of the bronchioles as shown in Figure 1. Pneumonia has several types including mycoplasma
pneumonia, viral pneumonia, bacterial pneumonia, and other types of pneumonia. Bacterial pneumonia occurs due to bacteria or fungi. Various symptoms are associated with the occurrence of bacterial pneumonia such as weakness of the body, old age, illness, poor nutrition, and weakened immune system. It is dangerous for people of all ages, but more dangerous for smokers, alcoholics, recent surgical patients, asthma, viral infection, and people having a frail immune system. Different viruses cause viral pneumonia such as flu and are accountable for almost 1/3 of all pneumonia cases. Mycoplasm pneumonia is also called atypical pneumonia and is caused by bacterium and generally affects all age people. Lobar pneumonia is one in which it usually affects one or more lobe/section out of the five lobes/section of the lungs (2 lobes in the left and 3 lobes at the right). Bronchopneumonia is one in which the pneumonia reaches the bronchial tubes. It is considered to be the most important and dangerous type of pneumonia all over the world, mostly found in children younger than 5 years, and causes death (approximately 12.9% of annual child deaths) [3, 4]. Pneumonia has several symptoms including fever, cough which produces mucus (greenish, yellowish, or bloody), squatness of breath, heavy sweating, tiredness, trembling, chest pain (which incurs with coughing and breathing), loss of appetite, turning color of lips and nails to blue, and confusion (in old age people). It is considered to be more dangerous for adults as well and is one of the leading causes of sickness and expiry across the world particularly in China [5–7].

In 2017, more than 850,000 people died from pneumonia. The death ratio due to pneumonia is very high in South Asia and Sub-Saharan Africa. According to a report published in 2017, the death ratio in five countries, i.e., Pakistan, India, Ethiopia, Nigeria, and the Republic of Congo, was more than half of the deaths from childhood pneumonia and was called the ultimate disease of poverty [8]. This shows that the mortality rate due to pneumonia has a strong correlation with the income of a country. In Japan, pneumonia is the third leading reason of expiry in elder people with age ≥80 years [9]. Approximately 1 million people are diagnosed with pneumonia disease and around 50k people die from this disease in a country like United States (US) every year. In Portugal, after lung cancer, pneumonia is the second most dangerous disease that leads to mortality due to respiration problems [10]. The list of the mortality rate due to pneumonia disease from 1990 to 2017 in all ages of people is shown in Figure 2. Pneumonia is a curable disease and does not spread from one country to another; its transmission is generally across local communities and can be controlled through basic health measures [11].

At the beginning of the 21st century, several coronaviruses have passed through the species fence to produce lethal pneumonia in human beings. In order to know the origin and development of these fatal pandemics, the experts need to inspect the structure of the virus and the method of how this virus causes infection. Furthermore, doing so will help the specialists in finding the right solution and providing proper treatment and possibly developing vaccines [13]. A short summary of the past epidemics and history of various types of coronavirus (MERS, SARS, and COVID-19) which have occurred over time are represented in Table 1.

Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) [14] is a serious breathing sickness problem that occurs due to coronavirus and has several indications like squatness of breath, temperature, coughing, and generally pneumonia. SARS appeared firstly in 2002 in China at Guangdong province and spread across the world. About 8 to 8.5k people got affected by this disease which results in 750–800 deaths [10, 15], with a fatal rate of about 10%. It is anticipated that this disease originated from bats [16]. The symptoms of SARS and flu are almost the same like headache, fever, chills, tiredness, and sometimes diarrhea.
After a few days, some other symptoms like higher temperature, fever, shortness of breath, and dry cough begin to appear as well [17].

Vhe Middle East Respiratory Syndrome Coronavirus (MERS-Cov) is a viral respiratory infection produced by a virus [18], which appeared for the first time in 2012 in the Middle East, Saudi Arabia [19, 20]. Some other cases of MERS disease were found in Jordan [21] and Qatar [22] and spread across the world. MERS is a zoonotic virus that was found mostly in camels and can be transmitted between humans and camels. According to the WHO report, humans are affected due to contact with the affected dromedary camels [23, 24]. MERS has various symptoms which include dizziness of breath, temperature, diarrhea, coughing, headache, vomiting, nausea, chest pain, and throat infection [22, 25, 26].

Nowadays, the world is facing a hazardous pandemic which occurred due to a virus and is named COVID-19, acknowledged in December 2019, for the first time in China, Wuhan Province, and lead to the death of many people [27–30]. COVID-19 is a type of coronaviruses which is found to be more dangerous and lethal than the other types [31]. The earlier cases of this disease were related to a seafood market in Wuhan, China, where live animals were kept for selling, and are considered to be a zoonotic origin of this pandemic [32]. The virus is transmitted from one person to another in three ways: (a) touching each other, (b) close contact (person-to-person), and (c) vaporizer transmission [33]. The most dangerous thing about COVID-19 is that it remains in incubation for up to two weeks (incubation period of 2 weeks) without any symptoms. COVID-19 has various symptoms such as shortness of breath, high temperature, fever, tiredness, pains, dry cough, sore throat, nausea, and flu, and some people will also have diarrhea [34]. Various techniques are investigated to identify this disease which include chest X-ray, CT scan, blood culture, sputum culture, fluid sample, bronchoscopy, and pulse oximetry.

Medical image analysis plays a vital character in the diagnosis of various ailments like MERS, Covid-19, pneumonia, etc., and is considered to be one of the auspicious approaches [35, 36]. Therefore, to detect pneumonia, chest X-ray images are used by various researchers. In addition, to analyze chest X-ray images accurately, there is a need for an expert radiologist who possesses expertise and experience in the desired domain. According to the World Health Organization (WHO) report, about 2/3 persons in the world still do not have access to the radiologist, in order to diagnose their disease. In order to overcome the issues mentioned above, this study proposes an intelligent computational framework based on ML and DL to detect
pneumonia disease in an efficient and effective manner. We used various Deep Convolutional Neural Network (DCNN) transfer learning techniques such as AlexNet, SqueezeNet, VGG16, VGG19, and Inception-V3 for extracting useful features from the image dataset. Six ML classifiers such as K-Nearest Neighbours (KNN), Logistic Regression (LR), Support Vector Machine (SVM), Naïve Bayes (NB), Ada-boost (AB), and Artificial Neural Network (ANN) have been investigated to diagnose whether a person has pneumonia or not. The proposed models have been trained and tested on chest X-ray and CT image dataset [37]. The performance of the proposed framework is tested on numerous performance measures such as accuracy, specificity, sensitivity, F1-measure, AUC-score, Mathew Correlation Coefficient (MCC), and ROC curve. It is expected that the suggested system will support the medical practitioners in order to diagnose pneumonia disease efficiently.

The remaining paper is organized as follows. Section 2 represents the review of the literature. The material and methods used in this study are discussed in Section 3. Section 4 demonstrates the results and discussion, and finally, we conclude our paper in Section 5.

2. Literature Review

Pneumonia is one of the fatal diseases, which is more dangerous for children and old age people. Toğaçar et al. [38] have used X-ray images of lungs for the identification of pneumonia. They have used CNN as a feature extractor by utilizing the existing models of CNN such as VGG-16 and AlexNet. These models extract a large number of features from images; for the reduction of the number of deep features, they used feature selection algorithms. Furthermore, they applied classical ML classifiers like, DT, LDA, and Linear regression for the diagnosis of pneumonia and achieved good results which show the importance of DL and classification algorithms. Liang and Zheng [39] have developed a framework based on DL for the diagnosis of child pneumonia using an image dataset and achieved satisfactory results. Jaiswal et al. [40] have proposed a DL-based method for the diagnosis of pneumonia using chest X-rays images. Their proposed classification/detection model was based on Mask-RCNN and achieved some good results which show the robustness and effectiveness of the model. Ge et al. [41] have investigated the prediction of pneumonia disease through ML (SVM, KNN, and DT) and DL (MLP and RNN) models and achieved promising results in terms of accuracy. Sirazitdinov et al. [42] have proposed an automated system for the forecasting of pneumonia on chest X-rays using ML algorithms. They used two types of CNN, i.e., Mask R-CNN and RetinaNet, and achieved satisfactory results.

Behzadi-Khormouji et al. [43] have presented a method based on DL and, specifically, CNN by using chest X-ray images and produced good results in terms of accuracy. To enhance the accuracy of the model, they used DCNN which was pretrained on the ImageNet data. In addition, they proposed a three-step preprocessing technique in order to enhance the generality of the model. Bhandary et al. [44] have proposed another healthcare framework based on DL for the diagnosis and detection of cancer and pneumonia. They have used two DL approaches, where the first one is a modified AlexNet. It was envisioned to separate and classify the chest X-rays (image dataset) into normal and abnormal classes by using SVM, and the performance of the proposed scheme was validated on pretrained DL transfer functions (VGG16 and AlexNet). On the other hand, the second approach implements a synthesis of handcrafted and the learned features in a person in order to increase the accuracy of lung cancer during valuation.

Medical imaging plays a significant part in the identification of numerous diseases [45, 46]. Classification of medical images is a significant and critical task to be accomplished. In order to classify the chest X-rays images and to diagnose pneumonia, this study depicts an extensive study of the fine-tuned versions of the latest Deep Convolutional Neural Network (DCNN) architectures (CNN, AlexNet, SqueezeNet, VGG16, VGG19, and Inception V3) for feature extraction and ML classification algorithms for the classification of pneumonia patients from a normal person.

3. Material and Methods

The following subsection describes the resources used and the approaches followed in carrying out this research study.

3.1. Dataset. The development of an automated and intelligent system extensively depends on the problem-related dataset. It means that a problem-specific dataset has a very high influence on the efficiency of an intelligent model. Considering the significance of the dataset, a chest X-ray and CT image dataset was used which is available online in the UCI Kaggle databases. The dataset consists of a total of 5856 images of two categories/classes, i.e., pneumonia and normal images. The dataset contains 1583 normal and 4273 pneumonia images. The dataset is distributed in two parts (training and testing), where 70% of the data is used to train the models while 30% of the data is used to test and validate the model. Figure 3 demonstrates an example of both categories/classes of chest X-ray images, where Figure 3(a) represents a normal image while Figure 3(b) represents the chest X-ray of a person having pneumonia.

3.2. Proposed System Methodology. The main goal and objectives of the proposed system are to diagnose whether a person has pneumonia or not at the early stages through their chest X-ray images in order to prevent them from more damage. In this study, the recent DCNN architectures based on the fine-tuned versions of (CNN, AlexNet, SqueezeNet, VGG16, VGG19, and Inception V3) are used to extract useful features from the images. Several preprocessing techniques are used, in order to present the data in a normalized form to the classification models. Various ML classification models such as KNN, SVM, LR, NB, AB, and ANN are used in this study. Different performance assessment metrics are computed to measure and track the performance of each utilized ML model. Keras deep learning framework is deployed which uses TensorFlow at the
backend for building and training our proposed system. The libraries and packages used in the implementation of the work include TensorFlow, Keras, Sklearn, Matplotlib, Seaborn, and NumPy. All the experiments were performed using the Jupyter NoteBook of the Anaconda integrated development environment (IDE). Figure 4 represents the framework of the proposed system.

3.3. Data Preprocessing. Data preprocessing is a vital technique used to provide data to the classification models in a well-organized manner, which are then trained and tested while using the normalized data. For the improvement of visual information quality (removal of noise, increasing contrast, deletion of high or low frequencies, etc.) of each input image, these images are preprocessed with the help of numerous techniques before being used in the classifiers. Preprocessing techniques such as intensity normalization, Contrast Limited Adaptive Histogram Equalization (CLAHE), and Min-Max normalization have been investigated in this study. Intensity normalization, CLAHE, and Min-Max normal distribution are interesting and important preprocessing techniques in image processing applications. Figure 5 represents the normal and images after applying the preprocessing techniques.

Looking at the dataset which represent two classes, i.e., pneumonia and normal images, almost 75% of the images represent pneumonia and the remaining 25% describe normal images which means that the dataset is imbalanced. To resolve the issue of unbalanced dataset and overfitting and to increase the accuracy of the models, various augmentation techniques have been used. The data augmentation techniques used include geometric transformations like rotations, zooms, rescale, shift, flips, and shears.

3.4. CNN Basic Architecture. CNN is a popular deep learning model particularly used for image classification problems. It normally consists of five layers which include the input layer, convolution layer, pooling layer, fully connected layer, and output layer. The practical assistance of CNN is having fewer parameters which significantly decrease the time it takes to learn and reduce the amount of data needed for training the model. In addition, CNN can be trained end-to-end for the extraction and selection of features from an image and, at last, can be used to predict or classify the images. It seems a bit difficult to know how a network understands or processes an image, but features conquered at various layers of a network perform better as compared to human-built features [47]. Figure 6 represents the basic architecture of the used CNN model.

The CNN architecture used for the experimental work in this study has the following properties:

1. Input layer: X-ray images are used as input and are provided at the input layer. The image dimensions are kept to 244 × 244.

2. Convolutional layer: we have used 3 convolution layers having 3 × 3 filter sizes and padding is set to zero.

3. Pooling layer: we have used max pooling for calculating the maximum value at every patch for each feature map. The max-pooling size is set to 2 × 2 while the stride value used is 2.0.

4. Fully connected layer: this layer used in the proposed architecture utilized the sigmoid activation function at the outer layer.

5. Output layer: the output layer gives us the predicted result that whether the person has pneumonia or not.

3.5. Deep Learning (DL) Architectures. DL architectures are extensively used in image processing specifically in healthcare for diagnosing various diseases. These DL techniques extract useful features from the images and present them to the models for further investigation. Here, in our study, we have used five important DL architectures such as AlexNet, VGG16, VGG19, Inception-V3, and SqueezeNet. A brief description of the investigated DL architectures is given hereinafter.

3.5.1. AlexNet. AlexNet is a type of CNN, which comprises various layers such as input, convolution, max pooling, dense, and output layers that are its basic building blocks. In 2012, it won the ILSVRC competition. It solves the problem
of image classification where the input image is one of 1000 different classes and the output is a vector of those classes. The $k$th element of the output vector is considered to be the likelihood that the input image belongs to the $k$th class. It may be noted that the sum of probabilities of the entire output vector is always equal to 1. AlexNet takes an RGB image as input having the size of 256 $\times$ 256, which means that all the images in the training and testing set need to have the size of 256 $\times$ 256. If the input image fails in matching the standard image size, then it needs to be converted to the standard size, i.e., 256 $\times$ 256 before using to train the network. If the input image used is a gray-scale image, then it is converted to RGB by replicating the single channel into a 3-channel RGB image. The architecture of AlexNet is changed from the CNN model which was used for computer vision problems and is much larger than CNN. AlexNet has 60 million parameters and 650,000 neurons which take a very long time for training.

3.5.2. VGG-16 and VGG-19. VGG (Visual Geometry Group) is a type of CNN architecture proposed for the first time by two researchers Simonyan and Zisserman in 2014 [48]. The VGG architecture won the ILSVR (ImageNet) competition in 2014. This architecture improves the AlexNet architecture by substituting the large kernel-sized filters, i.e.,
(11 in the first convolutional layer and 5 in the second convolutional layer) with multiple small 3 × 3 kernel-sized filters at the convolutional layer one after another and 2 × 2 in the max-pooling layer. At last, it has two fully connected layers followed by the activation function softmax/sigmoid for the output. The well-known VGG models are VGG-16 and VGG-19. The VGG-16 model consists of 16 layers while the VGG-19 model contains 19 layers. The main difference between both models is that VGG-19 contains one more layer at each of the three convolutional blocks.

3.5.3. Inception-V3. Inception models are a type of deep neural network (DNN) architecture developed by a researcher named Szegedy et al. for the first time in 2014, and the model was named as inception model [49]. The structures of inception models and the conventional CNN model are different from each other in such a way that inception models are inception blocks which means lapping the same input tensor with multiple filters and concatenating their results. There are various versions of the inception models. In 2015, Szegedy et al. [50] proposed a new version of the inception models named Inception-V3, which is an improved version of the previous versions of inception models, i.e., Inception-V1 and Inception-V2, and possesses more parameters. Inception-V3 contains a total of 24M parameters. The advancement in Inception-V3 was as follows: (a) it factorizes the "n × n" convolution into asymmetric convolutions, i.e., 1 × n and n × 1, (b) it factorizes the 5 × 5 convolutions into two 3 × 3 convolutions, and (c) it replaces 7 × 7 convolutions to a series of 3 × 3 convolutions. Actually, it consists of a block of convolutional layers which are arranged in a parallel manner and each layer consists of different sizes of filters 1 × 1, 3 × 3, and 5 × 5, respectively. Furthermore, 3 × 3 max pooling is also performed. The outputs are concatenated and sent to the next inception module.

3.5.4. SqueezeNet. SqueezeNet is a type of deep neural network developed by the researchers of Stanford University and was released on 22nd February 2016 for the first time. It is a type of CNN architecture consisting of 18 layers, particularly used in computer vision and image processing. The
3.6. Machine Learning (ML) Classification Algorithms. Various ML classification algorithms have been investigated for the diagnosis of whether a person has pneumonia disease or not. Each classification algorithm has its own importance, and its significance varies from application to application. In this paper, 6 distant natures of classification algorithms, namely, KNN, SVM, LR, NB, AB, and ANN, are applied in order to select the best and generalized prediction model.

3.7. Performance Measures. In order to track the performance of each classifier used in this study, several performance measures have been utilized such as accuracy, specificity, sensitivity, F1-measure, Mathew Correlation coefficient (MCC), AUC-score, and ROC curve. All the performance metrics are calculated by using the confusion table as shown in Table 2.

\[
\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \\
\text{specificity} = \frac{TN}{TN + FP} \times 100, \\
\text{sensitivity} = \frac{TP}{TP + FN} \times 100, \\
F1 - \text{measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \\
\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \\
\]

(1)

All of the abovementioned formulas are carried out from the confusion matrix which consists of the following basic components:

- True positive (TP): it means that the model prediction is positive and in actual fact the person has pneumonia. So, a pneumonia subject is diagnosed correctly by the model.
- True negative (TN): it means that the model prediction is negative and in actual fact the person does not have the pneumonia disease. Hence, a healthy person is diagnosed correctly by the classification model.
- False positive (FP): it means that the model did a wrong prediction by classifying a healthy person as a pneumonia patient. This is also known as type-1 error.
- False negative (FN): it means that the model did a wrong prediction by classifying a pneumonia patient as healthy. This is also known as type-2 error.

4. Results and Discussion

The simulation results of various ML classification algorithms by using different DL architectures such as AlexNet, SqueezeNet, VGG-16, VGG-19, and Inception-V3 are discussed in this section. These DL architectures, also called transfer learning techniques, extract useful features from the images which are very useful in classifying the normal and pneumonia patients in an efficient way. The performance of all utilized ML classifiers, i.e., KNN, SVM, LR, NB, AB, and ANN, was checked on the pneumonia chest X-ray dataset on full feature space generated by the transfer learning techniques. For measuring the performance of ML classifiers, different performance measures are used. In addition, preprocessing techniques are also applied to all features before being used by the classification algorithms.

4.1. Performance of Convolutional Neural Network (CNN) Classification Algorithm. This subsection represents the experimental results attained by the CNN classification algorithm. We performed multiple experiments on the basic CNN model by using various epoch numbers. First, we used 100 epochs and then 150 epochs, and at last, we used 200 epochs and got that the accuracy was increasing from epoch 0 to epoch 10, and after, that it becomes stable and remained 92.30%. Figure 7 represents the ROC curve obtained through the CNN classifier.

4.2. Performance of all Classifiers Using AlexNet Architecture. This section represents the simulation results carried out through all the utilized ML classification algorithms using the AlexNet transfer learning technique. Table 3 shows the experimental results attained through all the utilized 6 ML classification models.

Table 3 reveals that ANN outclassed all the other classifiers in terms of each performance measure. ANN attained the classification accuracy of 96.44%, specificity of 92.62%, and sensitivity of 96.82% as shown in Table 3. LR performed very well and achieved an accuracy of 95.94%, specificity of 91.40%, and sensitivity of 96.98% and stood second in
performances competition as shown in Table 3. SVM with kernel = “rbf” stood last in this regard as compared to other classification algorithms by achieving the classification accuracy of 51.13% as shown in Table 3.

Figure 8 shows the performance of all 6 ML classification algorithms using the features extracted through DL AlexNet architecture from the image dataset. From Figure 8, it is obvious that ANN outclasses other classification algorithms in terms of all performance measures where SVM with kernel = “rbf” performs poorly and stood last in the performance competition.

Figure 9 shows the F1-score and MCC score of all the utilized ML classification models using AlexNet DL architecture.

Figure 10 illustrates the ROC curves of all the utilized ML classifiers using AlexNet transfer learning architecture.

From Figures 9 and 10, it is observed that ANN produced good results and beats the rest of the classifiers in terms of all performance measures.

4.3. Performance of All Classifiers Using SqueezeNet Architecture. The experimental results and performances of all the utilized 6 ML classifiers, using the SqueezeNet transfer learning technique, are discussed here in this subsection. The transfer learning techniques are used to extract valuable features from the images and then present them to the classifiers to classify it. Table 4 shows the performances of all the 6 ML classification algorithms.

Table 4 demonstrates that ANN performs excellently in terms of all performance measures as compared to the rest of the classification models. ANN conquered the classification accuracy of 96.97%, specificity of 92.99%, and sensitivity of 97.52% as shown in Table 4. LR performs very well and achieved a classification accuracy of 96.24%, specificity of 92.99%, and sensitivity of 97.52% and stood second in performance competition as shown in Table 4. SVM with kernel = “linear” stood last in this regard as compared to other classifiers by achieving the accuracy of 52.03% as mentioned in Table 4.

Figure 11 signifies the performance of all utilized 6 ML classification algorithms using the features extracted through DL SqueezeNet architecture from the image dataset. From Figure 11, it is obvious that ANN outclasses other classification algorithms in terms of all performance measures where SVM with kernel = “rbf” performs poorly as compared to the rest of the classifiers.

Figure 12 shows the F1-measure and MCC score of all the utilized ML classification models using SqueezeNet transfer learning architecture. From Figure 12, it is observed that ANN beats all the other models in terms of F1-measure and MCC score by achieving the F1-score of 0.97 and MCC score of 0.92, respectively.

![Figure 7: ROC curve of the Convolutional Neural Network (CNN) classification algorithm.](image-url)
Figure 13 illustrates the ROC curves of all 6 ML classification algorithms on chest X-ray and CT image dataset using SqueezeNet transfer learning techniques.

4.4. Performance of All Classifiers Using VGG16 Architecture.
The experimental results and performances of all 6 ML classification algorithms using VGG-16 transfer learning techniques are described in this subsection. Table 5 notifies the performances of all 6 classifiers using VGG-16 architecture.

From Table 5, it is obvious that LR performs brilliantly in terms of the entire performance as compared to other classification models. LR attained an accuracy of 96.82%, sensitivity of 97.80%, and specificity of 94.03% as shown in Table 5. The second best model while using VGG16 architecture is ANN which attained an accuracy of 96.56%, specificity of 93.28%, and sensitivity of 97.52%. Again, SVM
Table 4: Performance of all classifiers using the SqueezeNet transfer learning architecture.

<table>
<thead>
<tr>
<th>Classification model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>F1-score</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN ($k = 5$)</td>
<td>95.16</td>
<td>96.44</td>
<td>91.05</td>
<td>97.80</td>
<td>0.95</td>
<td>0.73</td>
</tr>
<tr>
<td>SVM (rbf)</td>
<td>52.03</td>
<td>55.43</td>
<td>51.88</td>
<td>54.33</td>
<td>0.55</td>
<td>0.52</td>
</tr>
<tr>
<td>SVM (linear)</td>
<td>88.71</td>
<td>93.96</td>
<td>73.37</td>
<td>89.85</td>
<td>0.86</td>
<td>0.70</td>
</tr>
<tr>
<td>AB</td>
<td>90.12</td>
<td>93.13</td>
<td>81.43</td>
<td>87.34</td>
<td>0.90</td>
<td>0.74</td>
</tr>
<tr>
<td>NB</td>
<td>88.51</td>
<td>88.23</td>
<td>89.26</td>
<td>93.20</td>
<td>0.89</td>
<td>0.73</td>
</tr>
<tr>
<td>LR</td>
<td>96.24</td>
<td>97.62</td>
<td>91.94</td>
<td>99.20</td>
<td>0.96</td>
<td>0.90</td>
</tr>
<tr>
<td>ANN</td>
<td>96.97</td>
<td>97.52</td>
<td>92.99</td>
<td>99.40</td>
<td>0.97</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Figure 10: ROC curves of all 6 ML classifiers using the AlexNet transfer learning technique.

Figure 11: Performance of all classifiers using the SqueezeNet transfer learning architecture.
with kernel = “rbf” performed poorly as compared to other classification models and achieved an accuracy of 50.30% as shown in Table 5.

The performances of all 6 ML classification models, using VGG-16 transfer learning techniques, are shown in Figure 14.

Figure 15 shows the F1-score and MCC results of all 6 ML classification algorithms using VGG-16 transfer learning architecture while the ROC curves of all the utilized ML classification models are represented in Figure 16. It is observed from both Figures 15 and 16 that LR performed excellently in terms of these measures as compared to the other classifiers. The lowest performance was observed for SVM with kernel = “rbf” and stood last in this competition.


This subsection demonstrates the performance and experimental results obtained through all 6 ML classification models using the VGG-19 transfer learning technique. The transfer learning techniques extract useful features from the images and then present them to the classifiers for further processing. The performance of all 6 ML classification models is presented in Table 6.
Table 5: Performance of all classifiers using the VGG16 transfer learning techniques.

<table>
<thead>
<tr>
<th>Classification model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>F1-score</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN (k = 5)</td>
<td>95.41</td>
<td>96.00</td>
<td>93.66</td>
<td>98.32</td>
<td>0.95</td>
<td>0.88</td>
</tr>
<tr>
<td>SVM (linear)</td>
<td>81.72</td>
<td>81.23</td>
<td>82.99</td>
<td>87.33</td>
<td>0.83</td>
<td>0.59</td>
</tr>
<tr>
<td>SVM (rbf)</td>
<td>50.30</td>
<td>86.24</td>
<td>47.42</td>
<td>69.30</td>
<td>0.70</td>
<td>0.35</td>
</tr>
<tr>
<td>AB</td>
<td>90.12</td>
<td>93.57</td>
<td>80.01</td>
<td>86.82</td>
<td>0.90</td>
<td>0.74</td>
</tr>
<tr>
<td>NB</td>
<td>86.50</td>
<td>84.41</td>
<td>92.69</td>
<td>94.10</td>
<td>0.87</td>
<td>0.70</td>
</tr>
<tr>
<td>LR</td>
<td>96.82</td>
<td>97.80</td>
<td>94.03</td>
<td>99.51</td>
<td>0.97</td>
<td>0.92</td>
</tr>
<tr>
<td>ANN</td>
<td>96.56</td>
<td>97.52</td>
<td>93.28</td>
<td>99.22</td>
<td>0.97</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Figure 14: Performances of all classifiers using the VGG16 transfer learning technique.

Figure 15: F1-score and MCC of all classifiers using the VGG16 transfer learning technique.
Table 6 shows that the ANN classification model gives good performance while comparing it with the other models. ANN attained 97.01% classification accuracy, 97.62% sensitivity, and specificity of 93.80%. Sensitivity illustrates that the analytical test was positive and the person has pneumonia while the specificity shows that the diagnostic test was negative and the person does not have pneumonia and is healthy. LR also performed well and achieved good results, i.e., an accuracy of 96.92%, sensitivity of 97.60%, and specificity of 94.79%. Again, SVM with kernel = “rbf” shows the lowest performance by attaining the accuracy of 50.32%, specificity of 48.80%, and sensitivity of 84.30% as represented in Table 6.

Figure 17 demonstrates the performance of all 6 ML classification models using Inception-V3 architecture. ANN outclasses all the other classifiers in terms of accuracy, specificity, and specificity. SVM with kernel = “rbf” shows the lowest performance as shown in Figure 17. The MCC and F1-score results of all 6 ML classification algorithms using VGG-19 transfer learning architecture are described in Figure 18. From Figure 18, it is obvious that ANN performed brilliantly while SVM with kernel = “rbf” performed poorly.

Figure 19 demonstrates the ROC curves of all 6 ML classification algorithms on chest X-ray and CT image dataset using VGG-19 transfer learning techniques.

4.6. Performance of All Classifiers Using Inception-V3 Architecture. The performance and experimental results attained through all 6 ML classifiers using Inception-V3 DL architecture are discussed here in this subsection. The performance of all 6 ML classification models using Inception-V3 architecture is illustrated in Table 7.

Table 7 demonstrates that the ANN classification model performed really well in terms of the utilized performance measures as compared to the other models. ANN attained the classification accuracy of 97.19%, 97.88% sensitivity, and specificity of 94.92%. LR also showed good performance and achieved the accuracy of 97.08%, sensitivity of 97.90%, and specificity of 94.33%. Again, SVM with kernel = “rbf” performed poorly and remained at the last position in the classifier performance competition.

The MCC and F1-score of all classification algorithms using Inception-V3 DL architecture are described in Figure 21, while the ROC curves of all classification algorithms
using Inception-V3 DL architecture on chest X-ray images are demonstrated in Figure 22.

The performance of all five utilized DCNN transfer learning techniques and 6 ML classification algorithms have been evaluated by using various performance evaluation metrics as discussed above. From the abovementioned results, it is obvious that Inception-V3 and ANN performed brilliantly by attaining the classification accuracy of 97.19%, sensitivity of 97.92%, specificity of 94.92%, AUC of 99.53%, F1-measure of 0.97, and MCC of 0.92. LR along with Inception-V3 performs very well by gaining the classification accuracy of 97.08%, sensitivity of 97.90%, specificity of 94.33%, AUC of 99.52%, F1-measure of 0.97, and 0.92 of MCC score and stood second in this regard. The
Figure 19: ROC curves of all 6 ML classifiers using the VGG19 transfer learning technique.

Table 7: Performance of all classifiers using the Inception-V3 transfer learning techniques.

<table>
<thead>
<tr>
<th>Classification model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>F1-score</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN (k = 5)</td>
<td>94.20</td>
<td>94.24</td>
<td>94.03</td>
<td>97.80</td>
<td>0.94</td>
<td>0.85</td>
</tr>
<tr>
<td>SVM (linear)</td>
<td>85.02</td>
<td>92.03</td>
<td>86.41</td>
<td>87.33</td>
<td>0.85</td>
<td>0.78</td>
</tr>
<tr>
<td>SVM (rbf)</td>
<td>50.30</td>
<td>84.23</td>
<td>48.39</td>
<td>50.20</td>
<td>0.74</td>
<td>0.33</td>
</tr>
<tr>
<td>AB</td>
<td>87.40</td>
<td>91.33</td>
<td>78.91</td>
<td>83.62</td>
<td>0.87</td>
<td>0.67</td>
</tr>
<tr>
<td>NB</td>
<td>91.50</td>
<td>90.89</td>
<td>93.13</td>
<td>95.90</td>
<td>0.92</td>
<td>0.80</td>
</tr>
<tr>
<td>LR</td>
<td>97.08</td>
<td>97.90</td>
<td>94.33</td>
<td>99.52</td>
<td>0.97</td>
<td>0.92</td>
</tr>
<tr>
<td>ANN</td>
<td>97.19</td>
<td>97.88</td>
<td>94.92</td>
<td>99.53</td>
<td>0.97</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Figure 20: Performances of all classifiers using the Inception-V3 transfer learning technique.
Figure 21: $F_1$-score and MCC of all classifiers using the Inception-V3 transfer learning technique.

Figure 22: ROC curves of all 6 ML classifiers using the Inception-V3 transfer learning technique.

Table 8: A comparative study of the proposed system and previous approaches.

<table>
<thead>
<tr>
<th>Publications</th>
<th>Approach</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kermany et al. [51]</td>
<td>Convolutional neural network (CNN)</td>
<td>92.81</td>
</tr>
<tr>
<td>Stephen et al. [52]</td>
<td>DL model with 4 conv-layers and 2 dense layers</td>
<td>93.71</td>
</tr>
<tr>
<td>Saraiva et al. [53]</td>
<td>DL model with 6 conv-layers and 3 dense layers</td>
<td>95.29</td>
</tr>
<tr>
<td>Liang and Zheng [39]</td>
<td>DL model with 48 conv-layers and 2 dense layers</td>
<td>96.01</td>
</tr>
<tr>
<td>Wu et al. [54]</td>
<td>CNN + random forest</td>
<td>96.70</td>
</tr>
<tr>
<td>Proposed method</td>
<td>Intelligent framework (Inception-V3 + ANN)</td>
<td>97.19</td>
</tr>
</tbody>
</table>
performance of SVM with kernel = “rbf” is the worst among all the classification algorithms.

In addition, a comparative study of the proposed system is conducted with the previous ML and DL methods used in the past [39, 51–59]. A brief description of those approaches and the accuracies attained while using those approaches are demonstrated in Table 8.

Table 8 demonstrates a short summary of the previous approaches and classification accuracies attained via those techniques.

5. Conclusion

Pneumonia is an infective disease and is very hazardous for all ages and is more dangerous specifically for smokers, alcoholics, recent surgical patients, asthma patients, people with weakened immune systems, and children having an age of less than 5 years. The death ratio caused by pneumonia can be condensed if the patients are diagnosed at the initial stages and on-time medication and treatment is provided to them. This study proposes an ML- and DL-based intelligent predictive system for the diagnosis of pneumonia. Chest X-ray and CT image dataset was utilized for both training and testing of the proposed system. In order to improve the quality of visual information of each input image, various preprocessing methods such as intensity normalization, CLAHE, and Min-Max normalization have been utilized in this study. Five fine-tuned versions of DL transfer learning techniques such as AlexNet, SqueezeNet, VGG-16, VGG-19, and Inception-V3 were utilized to extract useful features from the X-ray images and then present them to the classifiers for further processing. Six imperative ML classification algorithms such as KNN, NB, ANN, SVM, LR, and AB were used to examine the efficiency of the proposed system. Numerous performance evaluation measures including classification accuracy, sensitivity, specificity, F1-score, AUC, MCC, and ROC were used to measure the performance of the proposed system. From the experimental results, it is observed that the Inception-V3 transfer learning technique and ANN performed brilliantly and attained the highest classification accuracy of 97.19%. Future work of this study includes the development of more optimized ML and DL algorithms that can significantly enhance the classification results. Further, developing an IoT-based real-time diagnosis of pneumonia disease is also one of the future works of this study.

Data Availability

All data is available in this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was supported by the Taif University Researchers Supporting Project Number TURSP-2020/126, Taif University, Taif, Saudi Arabia.

References


