Pneumonia Detection from Chest X-ray Images Based on Sequential Model

Asma Alshehri , Bayan Alharbi and Amirah Alharbi

Department of Computer Science, Umm Al-Qura University,
Makkah, Kingdom of Saudi Arabia

Abstract
Pneumonia is a form of acute respiratory infection that affects the lungs. According to the World Health Organization, pneumonia is the leading cause of death for children worldwide. As a result, pneumonia was the top killer of children under the age of five years old in 2015, which is 15% of all deaths worldwide. In this paper, we used CNN model architectures to compare between the result of proposed a CNN method with VGG based model architecture. The model's performance in detecting pneumonia shows that the proposed model based on VGG can classify normal and abnormal X-rays effectively and more accurately than the proposed model used in this paper.
Keywords:
X-ray images; pneumonia detection; image enhance

1. Introduction

Pneumonia is an infection of the lungs caused by bacteria, viruses, and fungi. In addition to viruses and bacteria, pneumonia can also be caused by fungi and flu complication. A person suffering from pneumonia may experience mild to severe symptoms. Depending on the cause of your pneumonia, the severity of your symptoms, and your age and overall health, you will need to receive the appropriate treatment. The average person recovers from pneumonia within a week or two, but it can be dangerous [1]. A flu shot can prevent pneumonia, and frequent hand washing and a pneumococcal pneumonia vaccine are offered to people at high risk. Lower respiratory tract infections, such as pneumonia, were the second leading cause of death in 2013, according to the Global Burden of Disease Study.

Pneumonia kills around 800,000 children under the age of five every year, with over 2200 deaths every day. Per 100,000 children, there are almost 1400 children infected with pneumonia [2]. The high incidence of pneumonia-related deaths among children has prompted scientists worldwide to develop more effective and rapid approaches for detecting pneumonia. As technology advances, more and more measurements are developed, the most common and beneficial of which are radiology-based procedures.

Chest X-ray, computed tomography (CT), and magnetic resonance imaging (MRI) are all diagnostic radiological techniques for lung diseases, and chest X-ray is the most effective. Diagnosing pneumonia using traditional methods is a difficult task even for the most skilled and experienced doctors, as X-ray images contain area information similar to other diseases, such as lung cancer. As a result, in this paper, we present a convolutional neural network to independently diagnose pneumonia using X-ray images, with an accuracy rate of 93 percent.

The paper is organized as follows. Section 2 presents literature reviews of medical image processing methods. An overview of Convolutional Neural Network (CNN) architecture is presented in Section 3 and illustrates a summary of the background of this machine learning and deep learning. In addition, it demonstrates the data used in this study, proposed methods, and training methods. We present the evaluation metrics in section 4. The results of the experiments are presented in Section 5. Finally, section 6 describes the conclusion of this study.

2. Related work

Several approaches are suggested by researches to detect lung diseases, some of which are described in this section. Several researches [3], [4], [5], [6], [7] suggested a variety of deep learning models for detecting lung cancer and other lung illnesses. The research in [3] focuses on detecting illnesses of the thorax. In [4], a 3D deep CNN with multiscale prediction techniques is presented for detecting lung nodules from segmented pictures. However, the work in [4] does not allow for the classification of disease categories, hence multiscale prediction methodologies are used for tiny nodules. In [5], a fully CNN is proposed for reducing the false positive rate in lung nodule classification. This method can only accepts CT scan images in order to limit the chances of making a false diagnosis.

In [5], the Luna 16 dataset is employed. In [6], a faster R–CNN was employed to detect the afflicted lung nodules while also lowering the FP rate. For object detection, a

Manuscript received April 5, 2022
Manuscript revised April 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.4.8

faster R–CNN yields promising results. For identifying and extracting nodule features, [7] uses a hybrid of deep CNN architecture and dual path network (DPN). In [8] uses a multi-patches arrangement with a Frangi filter to improve the performance of recognizing pulmonary nodules from lung X-ray images. Their approach, on the other hand, has a sensitivity of 94 percent and an FP rate of 15.1. The importance of artificial intelligence (AI) is demonstrated in [9] with a state-of-the-art in chest X-ray classification and analysis. Furthermore, the study [9] describes this problem in addition to structuring a new 108,948 front outlook database called ChestXray8, which contains 32,717 X-ray photos of distinct individuals. Deep CNNs are used by the authors in [9] to validate results on this lung data, resulting in encouraging results. In [1] A VGG-based CNN model is proposed to extract the feature of chest X-ray images detect Pneumonia. The proposed model consists of six layers, determined by cross validation to test the accuracy on the test set thus, applying six layers. First a 3×3 kernel convolution layers with a 1×1 strides after that, a 2×2 kernel max pooling strides are employed. In addition, they set some drop layers to randomly fit weights to zero. Next a two densely fully connected layers are used, followed by the Sigmoid function. The input channel is 224×224×1 whereas the output 0 denotes negative pneumonia case and 1 positive pneumonia case. The results of the training set exceeded 95% which indicate a great performance. In comparison to other models this proposed model represented the best results where the results are improved by each training epoch.

Table 1: Review of proposed work models.

No.	***Ref.***	Model and Approach
1	[3]	Weakly-supervised classification.
2	[4]	3D deep CNN model.
3	[5]	CNN model to reduce false positive rate.
4	[6]	Fast R–CNN model that lowers the FP rate.
5	[7]	Hybrid of deep CNN architecture and dual path network (DPN).
6	[8]	Multi patches arrangement with Frangi filter.
7	[9]	AI combined with Deep CNNs.
8	[1]	VGG-based CNN model.

3. Methodology

3.1 Data

Dataset was collected to evaluate model performance from the Kaggle competition; all chest X-ray images (anterior-posterior) were obtained from patients aged one to five years old. It consists of total 5786 X-ray images, which are organized into three folders (train, test, validation) and contain sub-folders for each image category (Pneumonia/Normal). Labels of the training subset where 0 denotes the images that do not show pneumonia, while 1 represents the images with pneumonia.

3.2 Data Pre-Processing

Preprocessing transforms raw data into an understandable format as part of a machine learning process. Deep learning processes employ image pre-processing techniques, which are not only popular and beneficial but also increase the quantity of the original dataset and enrich the implicit information in the dataset. In this study, we fixed the image size to 224×224 pixels and then explored the performance of different shapes and sizes, and the pixel scale was set to [0, 1] after dividing the maximum pixel of the image. Further details are illustrated in Table 2.

Table 2: Data pre-processing techniques used in this study.

Methods	***Setting***
Resize	224×224
Normalization	[(0, 255) → (0, 1)]
Zoom Range	0.3
width_shift_range	0.1
height_shift_range	0.1
Horizontal_Flip	True
vertical_flip	True

Pre-process chest X-ray images before inputting them into the proposed model. The Keras ImageDataGenerator class provides a host of different augmentation techniques like standardization, rotation, shifts, flips, brightness change, and many more. We are transforming an original ***image*** in different ways to produce multiple transformed copies of the same image. However, each copy differs in certain aspects depending on the used augmentation techniques. The results after applying image augmentation technique are shown in Figure 1.

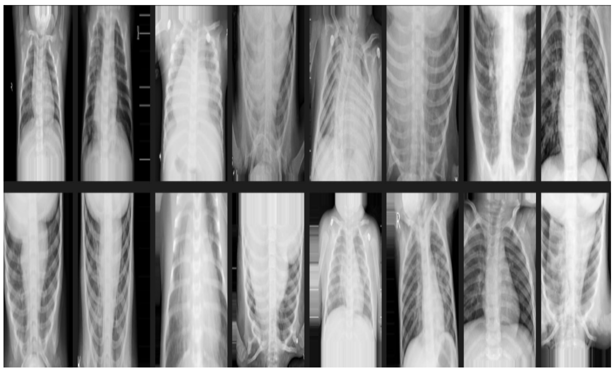

Figure 1. Examples of enhanced images.

In the chest radiograph, opacity determines whether a person has pneumonia or not. In case the opacity increase, then it would mean there is the occurrence of pneumonia.

From the above-displayed examples of X-Rays, we can observe that the difference between normal and pneumonia cases.

3.3 Proposed Network

Machine learning (ML) algorithms have gradually attracted the attention of researchers over the last few decades. Through given algorithms or specified steps, this type of algorithms could make use for image processing can be fully utilized.Nevertheless, traditional ML methods for classification require manual design algorithms or set feature extraction layers manually to classify images. As a result of the above situation, LeCun et al. [10]. proposed a CNN method Automated extraction of features through continuously stacked feature layers and output of the class that input images might belong to. The CNN model gradually extracts high-level features as the number of layers increases. Utilizing these advanced features, CNN examines the differences between different images and updates and records the learned parameters using a back-propagation algorithm.

The concept of CNN is to filter past images or feature maps through a specific convolution kernel to generate the feature map of the next layer and then combine this with methods like pooling to reduce feature map scale and computation. The resulting feature map is then given a nonlinear activation function to improve the model's characterization ability. Maximum and average pooling are two common pooling operations. The sole difference between the maximum and average pooling is the sub-region output, maximum pooling means that the feature provided to the pooling layer is divided into a number of sub-regions, with each sub-region outputting the maximum of its sub-regions based on horizontal and vertical strides, whereas average pooling outputs the average of each sub-region. ReLU (Rectified Linear Units) and Sigmoid are two common activation functions. ReLU has a derivative function and allows for backpropagation while simultaneously making it computationally efficient. Sigmoid function accepts any real value as input and outputs values in the range of 0 to 1. The image features are automatically extracted for the classification task by continuously stacking convolutional operations, pooling operations, nonlinear activation functions, and other fully-connected layers. Then, by evaluating these derived features, it is possible to determine whether the photos processed by the model exhibit pneumonia. The model's generalization ability is enhanced by making full use of the image's inherent pixel-level information.

3.4 Proposed CNN Model

A VGG-based CNN model is proposed by [1]to extract the feature of chest X-ray images to detect Pneumonia. The previous proposed model [1] consisted of six layers first a 3×3 kernel convolution layers with a 1×1 strides after that, a 2×2 kernel max pooling strides are employed. In addition, they set some drop layers to randomly fit weights to zero. Next a two densely fully connected layers are used, followed by the Sigmoid function. In the pre-processing step the Dynamic Histogram Enhancement (HE) technique is utilized to enhance the quality of the images. The HE technique modify the grey levels by the cumulative effort function thereby, enhancing the image quality. In this paper, we use a sequential based model architecture with higher layers than used in the original work. In addition, other techniques used in this paper to pre-process the images As illustrated above in Tabel 2. Figure 2 demonstrates the architecture of our proposed model, which was used to determine whether the input image contained pneumonia or not. The figure depicts our model, which has a total of seven layers and uses 3×3 kernel convolution layers with 1×1 stride and a ReLU activation function. A 2×2 strides kernel operation was used as a max-pooling operation after each convolution layer to keep the maximum of each sub-region, which is separated according to strides. In addition, a seventh layer is added to increase the model's performance by setting a drop layer to randomly fit weights to zero. Then, to take full advantage of the features collected by the preceding layers, two densely fully-connected layers are used, followed by the Sigmoid function, to output the possibility of patients suffering from pneumonia or not. As illustrated above, the input channel is 224 ×224 ×1 and the output size is 0 or 1, where 0 denotes that the image shows normal case, while 1 denotes that the image shows pneumonia.

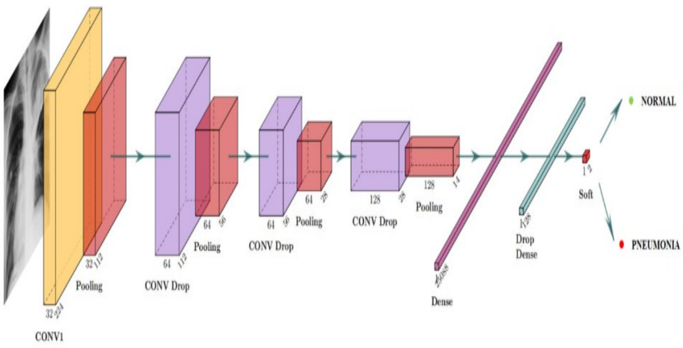

Figure 2. Details of proposed DL model.

4. Evaluation

4.1. Evaluation Metrics

To measure the performance of the model the following metrics are described accuracy, precision, recall and F1 score. The model's confusion matrix represents four indices, True Positive, True Negative, False Positive, False Negative; all of which are used to measure the performance of the model.

The true positive (TP) represents a "hit" a true pneumonia case predicted pneumonia case by the model. True negative (TN) represents "correct rejection" a negative pneumonia case predicted negative by the model. False positive (FP) represents -Type I error- "False alarm- a negative pneumonia case predicted positive by the model. False negative (FN) represents -Type II error- "miss" a positive pneumonia case predicated a negative by the model.

Model accuracy is defined as the number of classifications a model correctly predicts divided by the total number of predictions made. As shown in the formula.

$$Accuracey = \frac{(TN + TP)}{(TP + TN + FP + FN)} \tag{1}$$

Precision is the ability of a classification model to identify only the relevant data points and attempts to measure the actual positive classes. AS shown in the formula.

$$Precision = \frac{TP}{(TP + FP)} \tag{2}$$

Recall is the ability of a model to find all the relevant cases within a data set. It is also known as sensitivity true positive rate and hit rate. High recall value indicates most true positives are correctly classified. As shown in the formula.

$$Recall = \frac{TP}{(TP + FN)} \tag{3}$$

F-score is a way of combining the precision and recall of the model, and it is defined as the harmonic mean of the model's precision and recall. As shown in the formula.

$$F1 = \frac{Precision}{(Precision + recall} \tag{4}$$

Table 3 shows the fixed hyper-parameters. According to our architecture, The model is optimized using Adam optimizer, with a learning rate of 0.001. Adam is a deep learning replacement optimization model for training algorithm that replaces stochastic gradient descent. Adam combines the finest features of the AdaGrad and RMSProp methods to create an optimization technique for noisy issues with sparse gradients. Then, the Activation functions for hidden layers and last classification are set to ReLU and Sigmoid function used to fully exploit the information extracted from previous layers, determining whether or not individuals are suffering from pneumonia.

Table 3: Hyper-parameters of model.

Parameters	***Value***
Optimizer	Adam
Learning Rate	0.001
loss	binary_crossentropy
metrics	accuracy
Batch Size	16
Hidden Layer Activation Function	ReLU
Classification Activation Function	Sigmoid

5. Experiments and Results

5.1 Experiments

During the training procedure for 15 epochs with our proposed model, Figure 3 displays the training and validation accuracy and the training and validation loss obtained in Figure 4. The training accuracy was 93%, and validation accuracy was 63% also the training loss was 0.1777, and validation loss was 1.0549 , which means that the model we proposed has the potential power to acquire excellent performance. Furthermore, the training accuracy and training loss shows an apparent decreasing. As a result, the best model was obtained by develop a straightforward VGG based model architecture with fewer layers proposed by [1].

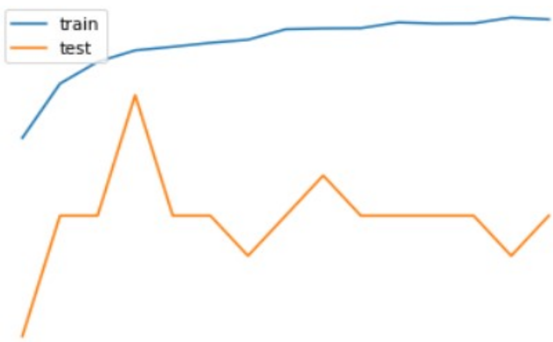

Figure 3. Training and validation accuracy.

5.2. Results

The difference in our model performance with VGG-based CNN model obtained by [1] illustrated in table 3. The Receiver Operating Characteristic Curve (ROC curve) was used to compute the predicted probability of CNN models using stacked convolution layers paired with pooling layers and ReLU activation or Sigmoid function, as illustrated in Figure 5. It reveals that just 8 samples out of 390 pneumonia images were misclassified, whereas 55 samples out of 234 actual normal images were predicted as pneumonia. A total of 63 images from the 624 images in the test subset were erroneously categorized. If you would like to itemize some parts of your manuscript, please make use of the specified style "itemize" from the drop-down menu of style categories.

Figure 4. Training and validation loss.

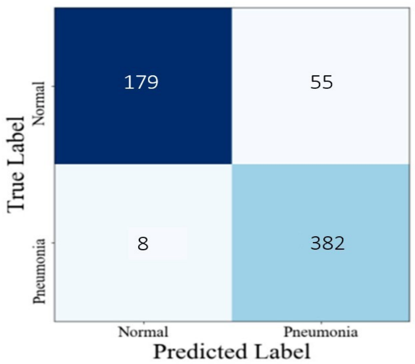

Figure 5. Confusion matrix of model.

Table 4: Summary of performance of proposed model and compared model

Model	Accuracy	Precision	Recall	F1 Score
Our model	93%	95.7%	76.49%	85.02%
Zhang et al. model [1]	96.24%	94.4%	90.8%	92.5%

6. Conclusion

A CNN-based model is applied to analyze chest X-ray images to diagnose pneumonia in this study. We used CNN model architectures to extract the features from original images, which contained seven layers combining ReLU activation function, drop operation, and max-pooling layers. Based on our proposed model accuracy rate of 93% and precision rate of 95.7%, we can conclude that it is accurate and performs well. In addition, we compared the results after training the last epoch between our proposed model with the VGG model. The best model was the VGG-based CNN model obtained by [1], due to several reasons i) using a pre-trained model, ii) training the model for 100 epoch, iii) utilizing HE technique in preprocessing images. We will continue to investigate more accurate classification architectures to diagnose more than two types of pneumonia in the future.

Acknowledgments
The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4310140DSR01).

References

[1] Zhang, D.; Ren, F.; Li, Y.; Na, L.; Ma, Y. Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network. Electronics 2021, 10, 1512.
[2] UNICEF DATA Percentage of deaths caused by pneumonia in children under 5 years of age (. https://data.unicef.org/topic/child-health/pneumonia/. 2021-7-20.
[3] Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2097–2106.
[4] Gu, Y.; Lu, X.; Yang, L.; Zhang, B.; Yu, D.; Zhao, Y.; Gao, L.; Wu, L.; Zhou, T. Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Computers in biology and medicine 2018, 103, 220–231.
[5] Setio, A.A.A.; Traverso, A.; De Bel, T.; Berens, M.S.; Van Den Bogaard, C.; Cerello, P.; Chen, H.; Dou, Q.; Fantacci, M.E.; Geurts, B.; others. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Medical image analysis 2017, 42, 1–13.
[6] Zhu, W.; Liu, C.; Fan, W.; Xie, X. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018, pp. 673–681.
[7] Kong, W.; Hong, J.; Jia, M.; Yao, J.; Cong, W.; Hu, H.; Zhang, H. YOLOv3-DPFIN: a dual-path feature fusion neural network for robust real-time sonar target detection. IEEE Sensors Journal 2019, 20, 3745–3756.

[8] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.
[9] Kallianos, K.; Mongan, J.; Antani, S.; Henry, T.; Taylor, A.; Abuya, J.; Kohli, M. How far have we come? Artificial intelligence for chest radiograph interpretation. Clinical radiology 2019, 74, 338–345.
[10] LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. Neural computation 1989, 1, 541–551.