The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200594

An Efficient Method to Predict Pneumonia from Chest X-Rays Using Deep Learning Approach

Uzair SHAH [a], Alaa ABD-ALRAZEQ [a], Tanvir ALAM [a], Mowafa HOUSEH [a] and Zubair SHAH [a,1]
[a] *College of Science and Engineering, Hamad Bin Khalifa University, Qatar*

Abstract. Pneumonia is a severe health problem causing millions of deaths every year. The aim of this study was to develop an advanced deep learning-based architecture to detect pneumonia using chest X-ray images. We utilized a convolutional neural network (CNN) based on VGG16 architecture consisting of 16 fully connected convolutional layers. A total of 5856 high-resolution frontal view chest X-ray images were used for training, validating, and testing the model. The model achieved an accuracy of 96.6%, sensitivity of 98.1%, specificity of 92.4%, precision of 97.2%, and a F1 Score of 97.6%. This indicates that the model has an excellent performance in classifying pneumonia cases and normal cases. We believe, the proposed model will reduce physician workload, expand the performance of pneumonia screening programs, and improve healthcare service.

Keywords. Deep learning, pneumonia detection, convolutional neural network

1. Introduction

Pneumonia is a severe health problem, common in individuals younger than five years and older than 65 years, causing millions of deaths every year [1]. Pneumonia diagnosis often begins with blood tests to identify infection, but precise identification is usually hard. Using chest X-rays is an alternative method for pneumonia diagnosis [2]. An early and accurate diagnosis from initial clinical inspection and interpretation of chest X-rays is always difficult, mostly due to the difficulty of detecting it and differentiating it from other possible pulmonary anomalies. Using chest radiography as part of the routine evaluation of diagnosing pneumonia patients is recommended [3]. In the interpretation of pneumonia from chest X-rays by a clinician, experience and capacity are important factors. However, subjectivity and discrepancies are often the key problems attributed to pneumonia diagnosis in reaching group consensus, making detection of pneumonia using chest X-rays a difficult task that needs skilled radiologists.

Diagnosing pulmonary diseases from chest X-rays have been investigated previously by many researchers [4-6] using the ChestX-ray14 dataset provided by the National Institute of Health with labels of pulmonary diseases. The dataset has more than 100,000 X-ray images with frontal views having annotations for 14 diseases. A dataset was extracted from the ChestX-ray14 dataset by Paul Mooney [7] which contains

[1] Corresponding Author, Zubair Shah, College of Science and Engineering, Hamad Bin Khalifa University, Qatar; Email: zshah@hbku.edu.qa.

pneumonia and normal cases. In this study, we used this pneumonia chest X-ray dataset and developed a deep learning architecture for the identification of pneumonia cases. The aim of this study was to develop an advanced deep learning-based architecture to precisely identify the absence or the presence of pneumonia symptoms in chest X-ray images and provide a diagnostic tool for pneumonia classification. The model we propose in this paper is based on the CNN architecture that is combined with various preprocessing techniques and parameters for fine-tuning.

2. Methods and materials

2.1. Study Data

For this study we considered pneumonia chest X-ray images, available from Kaggle [7], provided by Paul Mooney. The dataset contained 5856 chest X-ray images (normal= 1583, pneumonia=4273) with frontal view. The size of the images was 1024 x1024 pixels. We downscaled the size of the images to 224x224 using bi-cubic interpolation for inputting the images into the model. The dataset was split into 80% training and 20% testing. The training set was further split into 80% training and 20% validation. Hence, training set has 3747 images (normal=1014, pneumonia=2733), validation set had 937 images (normal=254, pneumonia=683), and the testing set had 1172 images (normal = 317, pneumonia = 855). The training and validation set was used for model training and validation and the performance of the trained model was evaluated using the test set. One-hot-encoding was used to convert the labels to categorical labels such as positive and negative. We evaluated the performance of the model using different metrics, such as accuracy, specificity, F1-score, precision, and recall.

2.2. Model Architecture

We propose a CNN model based on VGG16 architecture [8]. The model has 16-convolutional layers forming the feature extraction base for the network with two fully connected layers and a final logistic regression layer forming a predictive base for the network. A small 3x3 receptive field was used throughout the model that convolves with the input at each pixel (with stride 1). We also added a batch normalization layer after each convolutional layer for regularization. Batch normalization stabilizes the output of the convolutional layer [9]. We split the convolutional layers into four blocks. Every block consisted of two convolutional layers with batch normalization and one max-pooling layer that reduced the dimensionality of the feature maps to keep the important features. In the first two blocks, we used strides of 2x2 for pooling layers and for the last 2 blocks, we changed the strides to 3x3 to extract more specific features to generalize the network. After that, we used the flatten layer to take feature maps and convert them to a single vector for the fully connected layer. We added two fully connected layers to create the correct label prediction base for the network. We then used the SoftMax activation function to classify the labels at the final logistic regression layer. The model was trained using the pneumonia chest X-ray dataset. The weights of the networks were randomly initialized. The entire model was trained using the Adam optimizer using parameters (β_1=0.9, β_2=0.99) and the preliminary learning rate 1e^{-4}. The model was implemented using Python Keras library with Python TensorFlow library as a back-end and was trained on Kaggle freely available P100 GPUs.

3. Results

Table 1 shows the classification results for the test data in the form of a confusion matrix. Out of all test data (1172), the model correctly labelled 1132 images, which means that it had an excellent accuracy of 96.6% (Table 2). The model also correctly classified 839 of the 855 images belonging to pneumonia cases, meaning that it had high sensitivity/recall (98.1%). Of the 317 images belonging to normal cases, 293 images were labelled correctly by the model, indicating considerable specificity of 92.4%. Out of the 863 images classified as positive (i.e., pneumonia) by the model, the classification was correct in 839 images. This means that the model precision was high (97.2%) (Table 2). The model labelled 309 images as negative (i.e., normal), but the labelling was correct for 293 images. Thus, the negative predictive value (NPV) achieved by the model was 94.8%. Given that the model had high recall and precision, its F1 score was also high (97.6%). It is noteworthy that the accuracy of the model over validation and test sets were approximately equal (96.5% & 96.6%, respectively).

Table 1: Confusion matrix for the proposed model over the test data

		Predicted label		
		Pneumonia	**Normal**	**Total**
True label	**Pneumonia**	839	16	855
	Normal	24	293	317
	Total	863	309	1172

Table 2: Classification performance metrics for the proposed model over test data

Accuracy	**Sensitivity (recall)**	**Specificity**	**Precision (PPV)**	**NPV**	**F1 Score**
96.6%	98.1%	92.4%	97.2%	94.8%	97.6%

4. Discussion

In this study, we developed a deep learning model to discriminate between pneumonia cases and normal cases based on X-ray images. A total of 5856 high-resolution frontal view images of chest X-rays were used for the training, validating, and testing the model. The model achieved an accuracy of 96.6%, sensitivity of 98.1%, specificity of 92.4%, precision of 97.2%, NPV of 94.8%, and F1 Score of 97.6%. This indicates that the model has an excellent performance in classifying pneumonia cases and normal cases. The model proposed in this study showed superior performance in comparison to other existing models. To be more precise, Rajpurkar et al. [4] proposed a CNN-based model to identify pneumonia using chest X-rays obtained from ChestX-ray14 database. F1 Score was the only performance metric assessed in the study, and it was much lower than the F1 score results found in our model (43.5% vs. 97.6%) [4]. Another CNN-based model was used to detect pneumonia and other seven diseases based on chest X-rays obtained from the ChestX-ray8 database [5]. The model had a precision of 66%, recall of 93%, and F1 Score of 77%, indicating that it had lower performance than our model [5].

In the future, we plan to optimize our model by using more X-ray images from different sources to be universally applicable. We are also planning to extend the model to detect novel coronavirus disease (COVID-19) and other respiratory diseases. Furthermore, the performance of the model over X-rays images will be compared with its performance over computed tomography (CT) images. By demonstrating efficacy and accuracy of the model in detecting pneumonia and other diseases, we will upload the developed model in a free cloud-based platform to provide an instant diagnostic tool worldwide, especially in remote and economically underdeveloped areas where the access to skilled radiologists is limited. We recommend biomedical researchers to exploit our data and codes that will be uploaded to a publicly available database to improve the performance of the model and help move the field forward. Researchers should also assess the clinical and cost effectiveness of using such AI tools.

5. Conclusions

In this work, a deep learning model was developed to identify pneumonia using images of chest X-rays. The model showed an excellent classification performance, which was higher than the performance of models mentioned in previous studies. We expect that the model will reduce physician workload significantly, improve the performance of screening programs, improve healthcare delivery, and enhance the development of telemedicine. Further studies should be conducted to assess the effectiveness of this model. We will optimize the model by testing it using more images and extending it to detect other diseases (e.g., COVID-19) using CT images in addition to X-ray images.

References

[1] Dadonaite B, Roser M, Pneumonia, 2019.
[2] World Health Organization, Revised WHO classification and treatment of pneumonia in children at health facilities: evidence summaries, 2014.
[3] Niederman MS, Mandell LA, Anzueto A, Bass JB, Broughton WA, Campbell GD, Dean N, File T, Fine MJ, Gross PA, Martinez F, Marrie TJ, Plouffe JF, Ramirez J, Sarosi GA, Torres A, Wilson R, Yu VL, Guidelines for the Management of Adults with Community-acquired Pneumonia, American Journal of Respiratory and Critical Care Medicine, 163 (2001), 1730-1754.
[4] Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, Chexnet: Radiologist-level pneumonia detection on chest X-rays with deep learning, arxiv preprint arxiv, 1711.05225 (2017).
[5] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM, 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), 2097-2106.
[6] Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K, Learning to diagnose from scratch by exploiting dependencies among labels, arxiv preprint arxiv 1710.10501 (2017).
[7] Mooney P, Chest x-Ray Images (Pneumonia), 2018.
[8] Simonyan K, Zisserman A, Very deep convolutional networks for large-scale image recognition, arxiv preprint arxiv 1409.1556 (2014).
[9] Ioffe S, Szegedy C, Batch normalization: Accelerating deep network training by reducing internal covariate shift, arxiv preprint arxiv 1502.03167 (2015).