

# Deep Learning-Based Approach to Classify Pneumonia on Chest Radiographs

Aron Hernandez Trinidad<sup>1\*</sup>, Rafael Guzman Cabrera<sup>2</sup>, Blanca O Murillo Ortiz<sup>3</sup> and Teodoro Cordova Fraga<sup>1</sup>

<sup>1</sup>Physical & Engineering Department-DCI, University of Guanajuato campus Leon, Mexico

<sup>2</sup>Engineering Division, University of Guanajuato, Mexico

<sup>3</sup>IMSS High Specialty Medical Unit, T1 Clinic, Mexico

**\*Corresponding author:** Aron Hernandez-Trinidad, Physical & Engineering Department-DCI, University of Guanajuato campus Leon, Loma del Bosque 103, Lomas del Campestre, Leon, 37150, GTO, Mexico

## ARTICLE INFO

**Received:** 📅 November 02, 2024

**Published:** 📅 November 08, 2024

**Citation:** Aron Hernandez Trinidad, Rafael Guzman Cabrera, Blanca O Murillo Ortiz and Teodoro Cordova Fraga. Deep Learning-Based Approach to Classify Pneumonia on Chest Radiographs. Biomed J Sci & Tech Res 59(3)-2024. BJSTR. MS.ID.009298.

## ABSTRACT

The diagnosis of pneumonia by X-rays is a widespread practice in the early detection of the disease and initiation of timely treatment. Automatic classification models using convolutional neural networks (CNN) have proven to be an effective and accurate tool in the diagnosis of this disease. A model that detects pneumonia in a set of chest X-rays, using two CNNs: ResNet50 & VGG16 is proposed in this work. The method consists of preprocessing and classifying the set of radiographs available in the Kaggle repository that contains 2,224 records classified by an expert and segmented into two divisions: training set (1,600 images with 800 normal and 800 pneumonia) and test set (624 images with 234 normal and 390 pneumonia). This model achieves 97% accuracy by classifying the set of chest images into two categories: normal and pneumonia. The performance of the automatic classification scenario is robust and proves to be effective in the classification of chest X-rays to detect pneumonia.

**Keywords:** Pneumonia Diagnosis; Convolutional Neural Networks; Automatic Classification Model; Chest X-rays

**Abbreviations:** CNNs: Convolutional Neural Networks; HOG: Histograms of Oriented Gradients; SVM: Support Vector Machines; VGG: Visual Geometry Group; SVM: Support Vector Machines; TP: True Positives; FN: False Negatives; FP: False Positives; TN: True Negatives; OCT: Optical Coherence Tomography

## Introduction

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm used to analyze images. These networks are based on machine learning, where a computer learns to recognize patterns in the input data, in this case, chest X-ray images [1]. In pneumonia detection, CNNs can analyze specific image features that indicate the presence of disease, such as lung opacity, lung tissue consolidation, and inflammation [2]. In other words, CNNs are trained to identify specific patterns in images as edges, shapes, and textures [3]. By doing this, the network learns to extract relevant characteristics of interest from the image. The accuracy of CNNs in the detection of pneumonia on chest X-rays has been evaluated in several studies and has been shown to represent a solid auxiliary tool in the diagnosis of the disease [4-6]. They can analyze enormous amounts of images quickly and automatically, helping clinicians make informed decisions and

improve diagnostic efficiency. Although CNNs are powerful tools for image classification, they may still face some difficulties in doing so, including [7,8]:

- Variability:** Images can vary in size, quality, lighting, orientation, and other factors, which can make accurate classification difficult.
- Overlapping Classes:** In some cases, there can be similarities between different classes, which can make it difficult to differentiate them.
- Classification System:** The CNNs contain an integrated classification section that allows them to be evaluated, however, being integrated makes it difficult to modify parameters or replace the section with other advanced classification algorithms. In this paper, a model is proposed that solves the challenges to

generate a scenario for the automatic classification of chest X-rays to identify pneumonia optimally and accurately [9]. Instead of using CNNs as a complete and closed system, as is used in a wide range of computer vision and machine learning applications, they are used as feature extractors, allowing a model to be created consisting of three blocks: image preprocessing, feature extractor, and classification system [10]. This allows for taking advantage of each section in a better way, without affecting the complete system, to influence competitive classification values [11]. The main contributions of the work can be summarized as follows:

1. An automatic classification model is proposed to detect pneumonia in a set of chest radiograph images, using CNN as feature extractors.
2. The images go through a preprocessing that standardizes the set of radiographs with the same size (224X224) and filters: 2D convolution, low pass, high pass, Gaussian, median, and histograms of oriented gradients (HOG), which enhance the characteristics of lung opacity, suggesting the presence of pneumonia.
3. The dataset found in the Kaggle repository contains 2,224 chest X-rays. In addition, they are previously classified by experts and labeled into two classes: normal and pneumonia, which makes it possible to differentiate between both classes.
4. Two classification scenarios are proposed in the model: training/test set and crossvalidation that make up the classification system with support vector machines (SVM) as classifier algorithm and model evaluation metrics: Accuracy, Precision, F1 Score, and Recall.

The rest of this document is organized as follows. Section 2 presents relevant research and lists existing methods and models. In Section 3, the proposed model is explained in detail, including the image preprocessing, the CNNs used as feature extractors, and the classification system. In Section 4, the proposed model is evaluated, the advantages of our method are evaluated through experiments, and the results are reported. Section 5 concludes this work and forecasts our future work.

## Related Work

Detection of pneumonia using artificial intelligence is a constantly evolving area of research, and numerous studies have explored the use of the machine and deep learning techniques, using neural networks to improve the accuracy and effectiveness of detection of this disease. P. Rajpurkar, et al. [12] develop a model to detect pneumonia from

chest Xrays using a CNN. The authors evaluated the neural network's performance with 420 radiologist-labeled images, and it scored 95% on its evaluation metric, compared to the radiologists' 95% accuracy, demonstrating that the trained and labeled deep learning algorithm previously, achieve high precision in the detection of the disease. The authors suggested that using CNN models could potentially be used to assist radiologists in clinical settings, improving the speed and accuracy of pneumonia diagnosis. T. Rahman, et al. [4] present four different methodologies using four CNNs, through transfer learning. The pre-trained networks analyze a total of 5,247 chest images consisting of three labels: normal, viral, and bacterial. The accuracy of their model to classify the three labels was 93.3%, so the authors propose the model as a useful and rapid study for the diagnosis of pneumonia in the three corresponding classes. The computer-aided diagnostic tool can significantly help the radiologist to take more clinically useful images and identify pneumonia with its type immediately after acquisition. T. B. Chandra, et al. [13] expose a method for early and automatic detection of pneumonia in segmented lungs using machine learning.

The authors focus on the pixels of the region of interest that contribute to the extraction of relevant features of the confined area. The proposed model is examined with five classifiers on a data set of a total of 412 chest X-ray images divided into 206 normal cases and 206 with pneumonia, thus representing a binary classification model. The experimental results show that the method proposed by the authors achieves a significant precision of 95.63% with the best-modeled classifier, thus, it has surpassed existing conventional methods. These studies demonstrate the efficacy of CNNs and machine learning in detecting pneumonia on chest radiographs and their potential application in clinical settings.

## Proposed Method

A model consisting of three sections is proposed: preprocessing, feature vector, and classification system. Figure 1 illustrates this model. As can be seen, the data set is previously labeled into two classes: normal and pneumonia; We proceed to pre-process the size, quality, and lighting of the set using machine learning techniques. Once the first section has been completed, the images enter two pre-trained convolutional neural networks using transfer learning: ResNet50 and VGG16, both well-known for classifying chest images. However, the CNNs in the model work as extractors to obtain the feature vector from their convolutional bases, which represents an essential part of the method. By obtaining the vector of characteristics of the set of radiographs, the binary classification system is finally introduced to evaluate the method and obtain its performance.

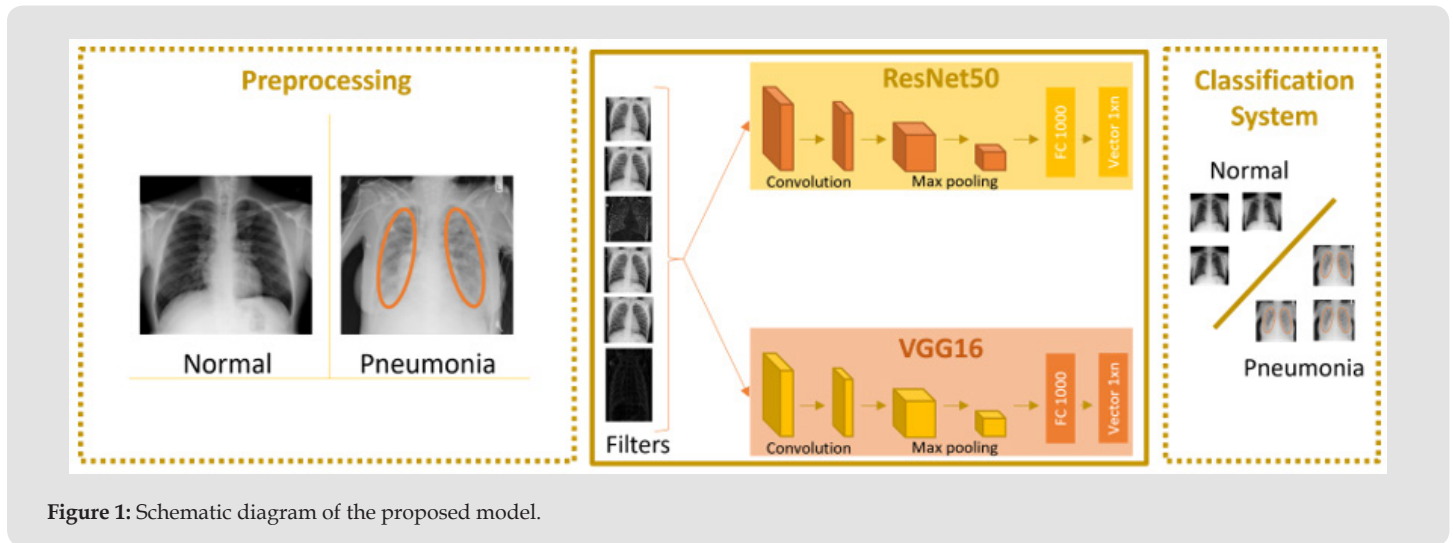


Figure 1: Schematic diagram of the proposed model.

### Dataset Preparation

Chest X-ray images used in this article were obtained from the Kaggle repository, a modified version of Paul Mooney [14], it is con-

taining 2,224 patient cases. The division of the data set can be seen in Figure 2.

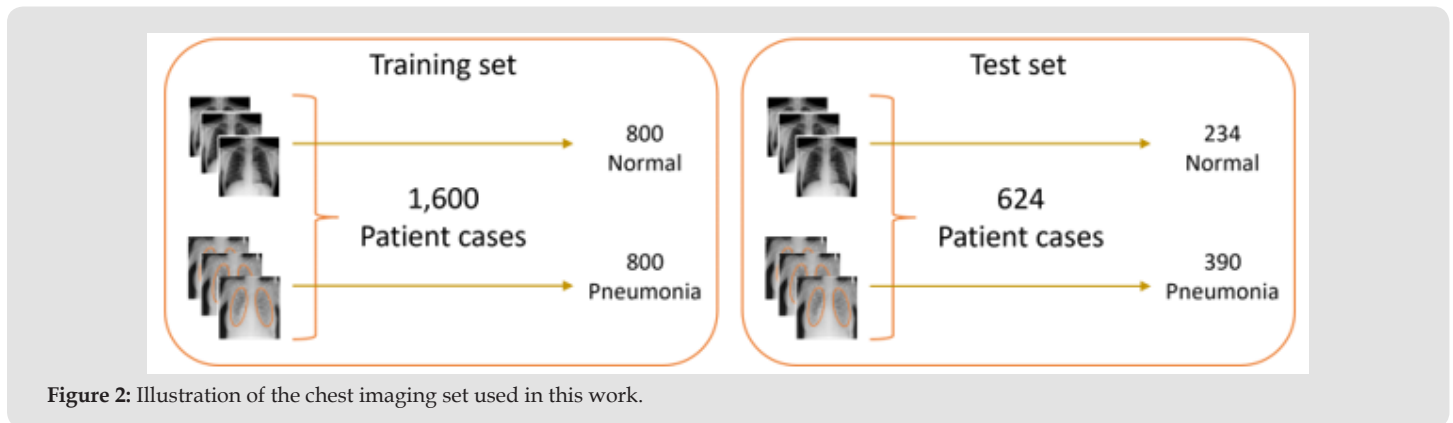


Figure 2: Illustration of the chest imaging set used in this work.

### Preprocessing

The performance, of the neural architectures, is affected by the quality of the input images, so six filters are proposed for the origi-

nal image to improve the characteristics of the set. In addition, ResNet50 and VGG16 are pre-trained networks with an input dimension of 224×224, so the set of images is resized to that size. Figure 3 shows the corresponding preprocessing.

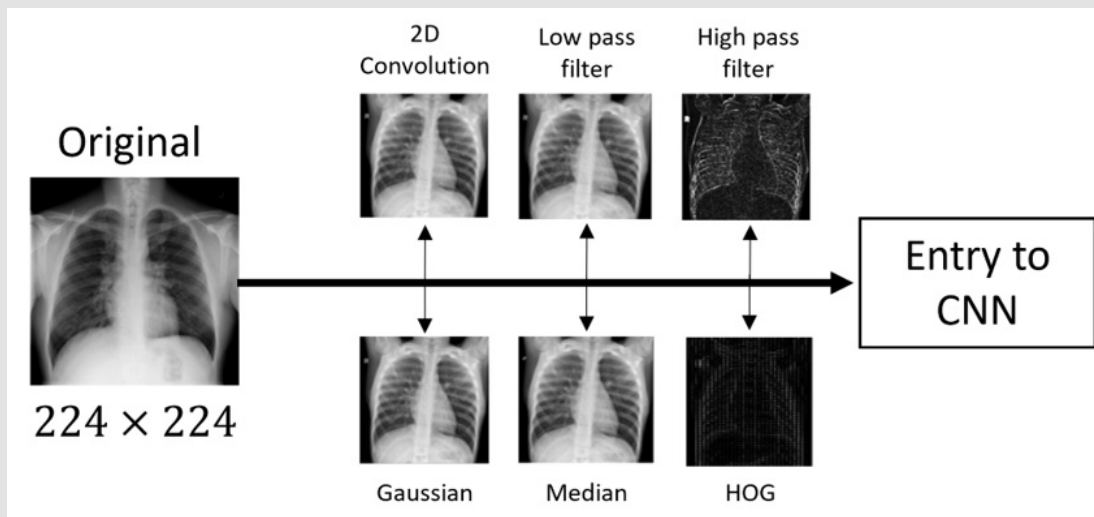


Figure 3: Results of the preprocessing of the original image whose output is the input to the neural architectures.

### Convolutional Neural Networks

**ResNet50:** ResNet50 is a deep convolutional neural network model used for object recognition and image classification tasks. It was developed by Microsoft researchers in 2015 and has 50 layers [15]. The name “ResNet” refers to the residual connections used in its architecture (Figure 4). These connections allow data to move directly from one layer to another, avoiding gradient fading problems

and improving the ability of the network to learn and generalize from training data. ResNet50 is widely used in machine vision applications and has demonstrated high performance in image classification tasks on large and complex data sets [16,17]. The notation (k k, n) in each convolutional layer (conv) block denotes the size of the square matrix k and n channels, below each stage represents the repetition of each unit, the penultimate layer FC 1000 is the fully connected layer with a thousand neurons. Finally, 1×n represents the feature vector.

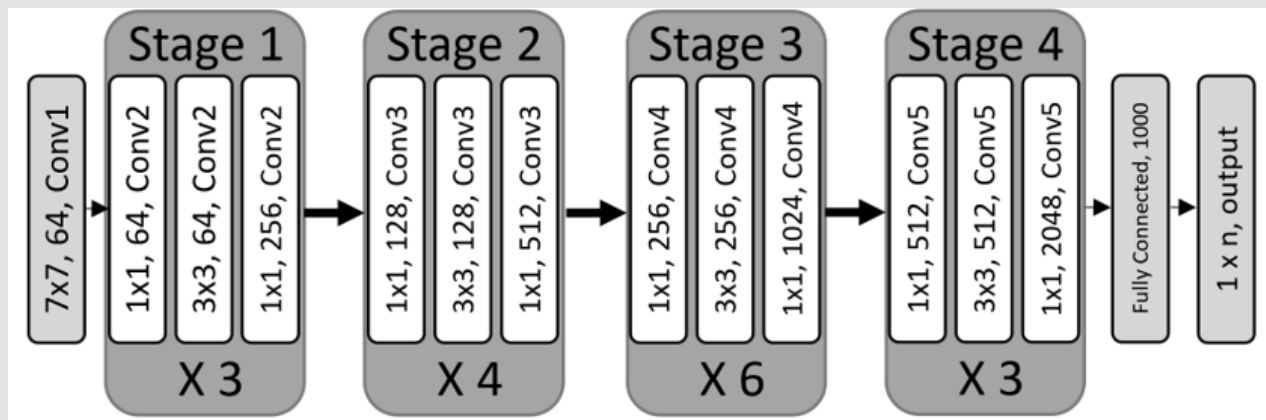


Figure 4: Structure of the ResNet50 network.

**VGG16:** VGG16 is a convolutional neural network model used for image classification tasks. It was developed by researchers at the Visual Geometry Group (VGG) at the University of Oxford in 2014 and has 16 layers [18]. The VGG16 architecture is characterized by having very deep convolutional layers and small 3×3 convolution filters, which allows better feature extraction in high-resolution images

(Figure 5). The model is trained on large and complex data sets, such as ImageNet, and has shown high performance in image classification tasks [19]. It is widely used in machine vision applications and has served as the basis for the development of other convolutional neural network models [20]. The input of the first convolution layer (Conv2D) is an image of size 224X224. For all convolution layers,

the convolution kernel is 3X3 in size. The layers are accompanied by Max-Pooling layers (MaxPool) that reduce the size of the filters during training. At the output of the convolution and pooling layers, there are

three layers of fully connected neurons: Dense, Dense, and softmax to determine the image class.

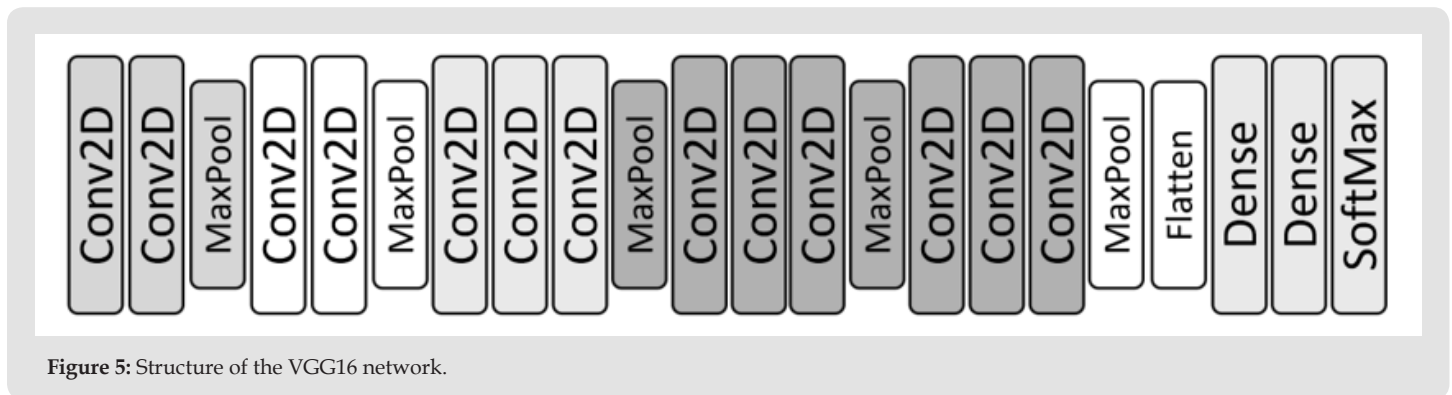


Figure 5: Structure of the VGG16 network.

## Classification System

Cross-validation and the train/test set are two common techniques for evaluating the performance of a machine learning model. Both methods are used to assess a model's ability to generalize to unseen data, but they differ in the way the data is partitioned for evaluation [21]. Cross-validation is a method that involves dividing the data set into  $k$  folds. The model is trained  $k$  times, each time using  $k-1$  folds for training and the remaining fold for evaluation. Performance scores are calculated for each fold and averaged to get an overall measure of model performance. Cross-validation helps to reduce model evaluation scarcity and to obtain a more accurate estimate of model performance. The training/test set, on the other hand, involves dividing the data set into two exclusive sets: a training set and a test set. The model is trained on the training set and evaluated on the test set. The performance score is calculated based on the model's ability to generalize data not seen in the test set. In general, crossvalidation is used when you have a limited data set and you want a more accurate estimate of model performance, while a train/test set is a faster and easier technique used when you have a large data set, and a rapid evaluation of the model is desired. From 2,224 images in the X-ray set, the training and testing scenario was divided into 80% and 20%, respectively. For the case of cross-validation,  $k = 20$  folds were performed on the entire data set. The machine learning algorithm, Support Vector Machines (SVM) was used for the binary classification of the ensemble. The SVM classifier has had favorable and relevant results to separate two data classes using a hyperplane in high-dimensional feature space [22,23].

## Confusion Matrix and Evaluation Metrics

A confusion matrix is an evaluation tool used in machine learning and statistics to assess the accuracy of a classification model. The confusion matrix shows the number of times the test samples are classified correctly and incorrectly, relative to the true class labels. The confusion matrix is a  $2 \times 2$  table showing four possible outcomes of the classification [24]:

- **True Positives (TP):** The model correctly classified the positive samples.
- **False Negatives (FN):** The model incorrectly classified the positive samples.
- **False Positives (FP):** The model incorrectly classified the negative samples.
- **True Negatives (TN):** The model correctly classified the negative samples.

From these results, the model evaluation metrics are calculated: Accuracy, sensitivity, specificity, and the F1 score [25]. Accuracy is defined as the number of true predictions divided by the total number of predictions made. Sensitivity (also known as the true positive rate) is defined as the number of true positives divided by the total number of true positives and false negatives. Specificity (also known as the true negative rate) is defined as the number of true negatives divided by the total number of true negatives and false positives. The F1 score is a measure that combines precision and sensitivity in a single measure. The confusion matrix is an especially useful tool for evaluating the quality of the classification model and provides valuable information about the performance of the model in terms of its ability to correctly classify the test samples into the true classes.

## Results

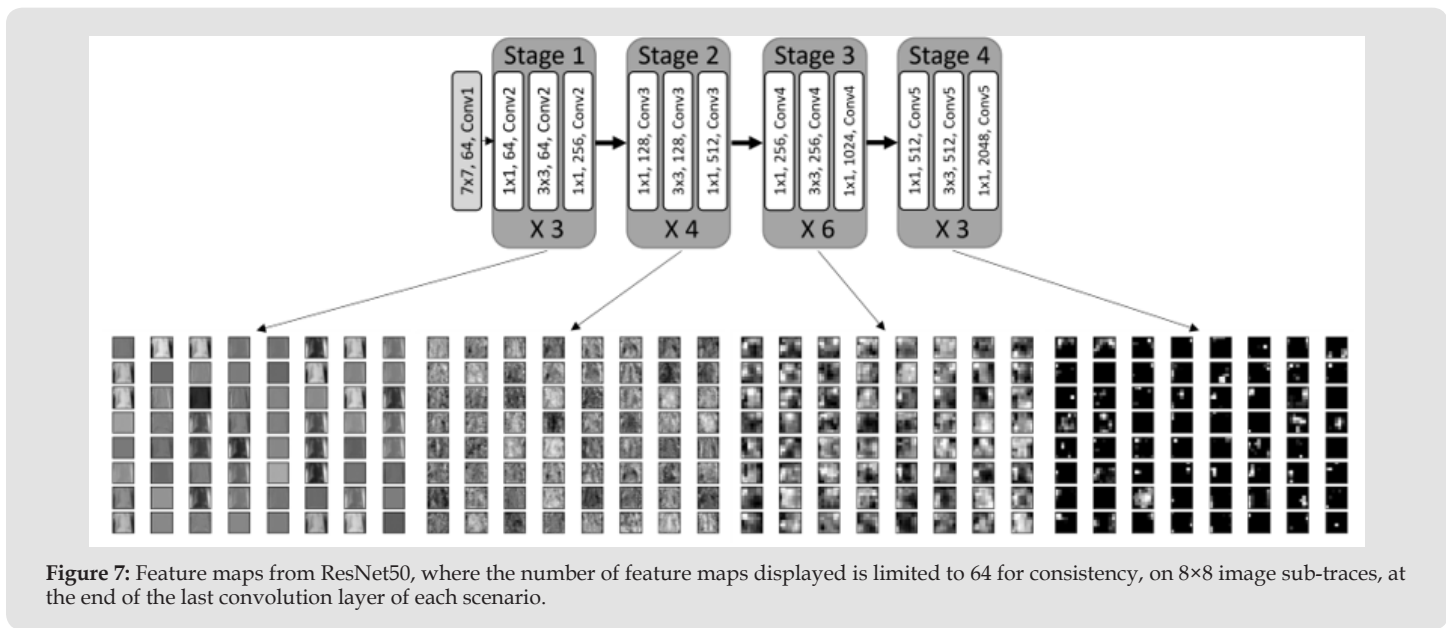
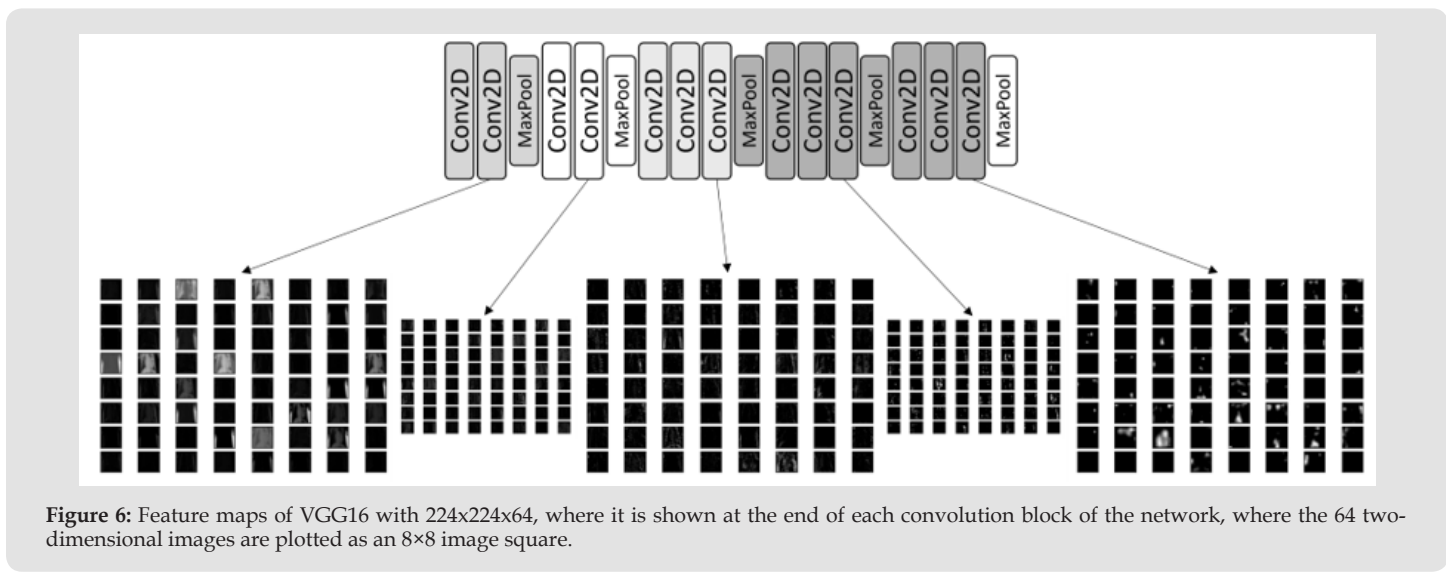
Normal chest X-ray images and those with the presence of pneumonia were preprocessed as shown in Figure 3. These, in turn, were the input to the CNNs, pre-trained deep learning neural networks are opaque, which means that, although they can make useful and skillful predictions, it is not clear how or why a given prediction was made. In our model, the CNNs function as feature extractors from their convolutional basis, to obtain the feature vector that is passed to the classification system discussed above. However, it can inspect the two-dimensional filters of the convolutional layers to understand exactly which features were detected from a given X-ray input.



### Characteristic Maps of a Chest X-ray with the Presence of Pneumonia

ResNet50 and VGG16 architectures are part of the proposed model, which are a fundamental part of the prediction and detection of pneumonia in a set of chest images. Both deep neural networks are constituted by convolutional blocks (convolutional bases) so feature maps were generated at the end of each block, to inspect the detection characteristics of the CNNs. In Figures 6 & 7, the results are shown. The result, of applying the filters, in the first convolutional layer of each neural network is many versions of the chest X-ray image with the presence of pneumonia with different highlighted features, for example, some highlight lines, while others sharpen the background or

the foreground. The feature maps closest to the input of the network architecture capture a great deal of fine detail from the original image, as the network progresses the feature maps show less and less detail. This pattern was to be expected since features are abstracted into more general concepts that are used to perform the classification. Although the x-ray image is not clear in the final block, the ability to interpret these deeper feature maps is lost, but they are an essential part of generating the feature vector. From the previous result that was applied to the original image of the two available classes: normal and pneumonia, we proceeded to obtain the feature vectors for the filters, in the two classification scenarios and the binary classifier, Tables 1 & 2 show the values obtained from the proposed model with the corresponding considerations (pages 10 and 11).



**Table 1:** Results of the evaluation of the model with ResNet50.

Images	Filter	Training/Test Set Scenario				Cross-validation Scenario			
		Accuracy	Sensitivity	Specificity	F1 Score	Accuracy	Sensitivity	Specificity	F1 Score
	Original	0.91	0.91	0.9	0.9	0.95	0.95	0.95	0.95
	Conv2D	0.92	0.91	0.9	0.91	0.95	0.95	0.95	0.95
	High pass	0.63	0.71	0.69	0.62	0.93	0.93	0.93	0.93
2,224	Low pass	0.92	0.92	0.9	0.91	0.96	0.96	0.95	0.95
	Gaussian	0.91	0.91	0.9	0.9	0.95	0.95	0.95	0.95
	Median	0.89	0.9	0.86	0.87	0.94	0.94	0.94	0.94
	HOG	0.88	0.87	0.86	0.86	0.91	0.91	0.91	0.91

**Table 2:** Results of the evaluation of the model with VGG16.

Images	Filter	Training/Test Set Scenario				Cross-validation Scenario			
		Accuracy	Sensitivity	Specificity	F1 Score	Accuracy	Sensitivity	Specificity	F1 Score
	Original	0.93	0.93	0.92	0.93	0.94	0.94	0.94	0.94
	Conv2D	0.93	0.94	0.92	0.93	0.94	0.94	0.94	0.94
	High pass	0.92	0.93	0.9	0.91	0.96	0.96	0.96	0.96
2,224	Low pass	0.95	0.94	0.94	0.94	0.97	0.97	0.97	0.97
	Gaussian	0.92	0.92	0.9	0.91	0.94	0.94	0.94	0.94
	Median	0.92	0.92	0.9	0.91	0.94	0.94	0.94	0.94
	HOG	0.92	0.92	0.9	0.91	0.94	0.94	0.94	0.94

## Discussion

The proposed model obtained optimal and competitive values with the state of the art. Image preprocessing acquires an important relevance to generate an efficient and accurate evaluation that manages to classify the classes with 97% in the best classification scenario. From the results obtained from the previous tables, the model improves its evaluation in all filters in the training/test set classification technique with the convolutional basis of VGG16, with its counterpart ResNet50 in the same technique. While in the cross-validation technique, it is the other way around, ResNet50 slightly outperforms VGG16 in most filters. However, it is observed that by applying the low-pass filter to the data set, it is possible to obtain the highest values in both techniques and architectures, respectively. The model generates the best evaluation with 95% accuracy in the training/test set and 97% in the same evaluation metric in cross-validation with the VGG16 convolutional base. Although both architectures are popular in computer vision and are used for similar tasks with differing architectural depth, the VGG16 network has slightly significant performance for classifying and detecting pneumonia in a set with two classes: normal and pneumonia, from chest radiographs, applying preprocessing to the images and classifying with SVM in the classification system. The results show that there is a high-reliability index in the proposed model.

## Conclusion

The evaluation of the automatic pneumonia classification model from a set of chest X-rays was optimal, precise, competitive, and efficient, having values of 95% accuracy in the training/test set classification scenario with preprocessing and low pass filter and 97% under the same conditions in the cross-validation scenario, with the VGG16 convolutional neural network. The choice of a deep learning model for a specific task depends on several factors, including the size and complexity of the data set, the availability of computational resources, the experience of the developer, and the model's performance on different metrics. In the specific case of pneumonia classification on chest radiographs, it has been shown that VGG16 may perform better than ResNet50, despite being a shallower model. This is partly because VGG16 was designed specifically for image classification, while ResNet50 was originally designed for more general object detection and classification tasks. Furthermore, VGG16 has a simpler and more uniform architecture than ResNet50, which can make it easier to train and optimize. It is also possible that the data set used to train the models is not large or complex enough to take full advantage of the additional depth of ResNet50. In conclusion, the choice of a deep learning model depends on several factors, and it is not always necessary to use the deepest model to achieve the best performance on a specific task. It is important to evaluate different models and select the one that best suits the specific needs of the project.

## Acknowledgments

The authors thank the University of Guanajuato for the partial support in the development of this work under the project DAIP/2023-59023.

## Conflict of Interest

The authors declare that they have no conflict of interest that could influence the results, or the interpretation of the findings presented in this article. Furthermore, we confirm that the work presented here is original and has not been previously published elsewhere.

## Data Availability Statement

The data that support the findings of this study are openly available in Kaggle at <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>, from: Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, V2, doi: 10.17632/rschjbr9sj.2

## References

- Li Z, Liu F, Yang W, Peng S, Zhou J (2021) A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems* 33: 12.
- Goyal S, Singh R (2021) Detection and classification of lung diseases for pneumonia and covid-19 using machine and deep learning techniques. *Journal of Ambient Intelligence and Humanized Computing* 14(4): 3239-3259.
- Geirhos R, Rubisch P, Michaelis C, Bethge M, Wichmann F A, et al. (2018) Imagenet-trained cnns are biased towards texture; increasing shape bias improves accuracy and robustness. *arXiv preprint*.
- Rahman T, Chowdhury M E, Khandakar A, Islam K R, Islam K F, et al. (2020) Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. *Applied Sciences* 10(9): 3233.
- Saraiva A A, Ferreira N M F, Sousa L L, Costa N J C, Sousa J V M, et al. (2019) Classification of images of childhood pneumonia using convolutional neural networks. In: *Bioimaging*, pp. 112-119.
- Stephen O, Sain M, Maduh U J, Jeong D U (2019) An efficient deep learning approach to pneumonia classification in healthcare. *Journal of healthcare engineering*.
- Ruthotto L, Haber E (2020) Deep neural networks motivated by partial differential equations. *Journal of Mathematical Imaging and Vision* 62: 352-364.
- Shafiq M, Gu Z (2022) Deep residual learning for image recognition: A survey. *Applied Sciences* 12(18): 8972.
- Mamalakis M, Swift AJ, Vorselaars B, Ray S, Weeks S, et al. (2021) Denres-cov-19: A deep transfer learning net- work for robust automatic classification of covid-19, pneumonia, and tuberculosis from x-rays. *Computerized Medical Imaging and Graphics* 94: 102008.
- Jia G, Lam H K, Xu Y (2021) Classification of covid-19 chest x-ray and ct images using a type of dynamic cnn modification method. *Computers in biology and medicine* 134: 104425.
- Zhao X, Qi S, Zhang B, Ma H, Qian W, et al. (2019) Deep cnn models for pulmonary nodule classification: model modification, model integration, and transfer learning. *Journal of X-ray Science and Technology* 27(4): 615-629.
- Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, et al. (2017) CheXnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv*.
- Chandra T B, Verma K (2020) Pneumonia detection on chest x-ray using machine learning paradigm. In: *Proceedings of 3rd International Conference on Computer Vision and Image Processing: CVIP 2018* 1: 21-33.
- Kermany D S, Goldbaum M, Cai W, Valentim C C, Liang H, et al. (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. *cell* 172(5): 1122-1131.
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778.
- Theckedath D, Sedamkar R (2020) Detecting affect states using vgg16, resnet50 and se-resnet50 networks. *SN Computer Science* 1: 1-7.
- Mukti I Z, Biswas D (2019) Transfer learning based plant diseases detection using resnet50. In: *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*, p. 1-6.
- Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv*.
- Krishnaswamy Rangarajan A, Purushothaman R (2020) Disease classification in eggplant using pre-trained vgg16 and msvm. *Scientific reports* 10(1): 2322.
- Qassim H, Verma A, Feinzimer D (2018) Compressed residual-vgg16 cnn model for big data places image recognition. In: *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 169-175.
- Afendras G, Markatou M (2019) Optimality of training/test size and resampling effectiveness in cross-validation. *Journal of Statistical Planning and Inference* 199: 286-301.
- Yang Y, Li J, Yang Y (2015) The research of the fast svm classifier method. In: *2015 12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 121-124.
- Kumari R, Srivastava S K (2017) Machine learning: A review on binary classification. *International Journal of Computer Applications* 160(7).
- Tarca A L, Carey V J, Chen X w, Romero R, Draghici S (2007) Machine learning and its applications to biology. *PLoS computational biology* 3(6): e116.
- Hossin M, Sulaiman M N (2015) A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process* 5(2): 1-11.



ISSN: 2574-1241

DOI: 10.26717/BJSTR.2024.59.009298

Aron Hernandez-Trinidad. Biomed J Sci & Tech Res



This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>



#### Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>