STM JOURNALS

Journal of
Microwave Engineering & Technologies

ISSN: 2349-9001
Volume 11, Issue 3, 2024
September–December
DOI (Journal): 10.37591/JoMET

https://journals.stmjournals.com/jomet

Review JoMET

Pneumonia Detection and Classification Using Deep Learning

Saniya P.M.[1,*], Mohammad Shabeer[2], Muhammed Sinan M.S.[2], Muzammil Rahman K.M.[2], Sanad Fazal Kota[2]

Abstract

Pneumonia, an infectious lung disease primarily caused by bacteria, often exacerbated by environmental factors, leads to the accumulation of pus in the lung's alveoli. Accurate diagnosis through chest X-rays, ultrasounds, or lung biopsies is crucial to avoid misdiagnosis and ensure proper treatment, crucial for patients' quality of life. Diagnostic capacities have been greatly improved by deep learning advances, especially with convolutional neural networks (CNNs). This research presents a robust CNN-based approach for predicting and detecting pneumonia from chest X-ray images. Using a dataset comprising 20,000 images at a resolution of 224x224 and trained with a batch size of 32, the CNN model achieved an impressive 95% accuracy during training. The study demonstrates the CNN model's effectiveness in identifying COVID-19, bacterial, and viral pneumonia solely from chest X-ray images, highlighting its potential for accurate diagnosis in clinical settings. It was widely believed at the time that computers would eventually become as adaptive as humans. The fundamental idea behind adaptive learning is that the system or tool may adapt to the user's or student's preferred learning style, giving them a better and more efficient learning experience.

Keywords: Pneumonia detection, adaptive deep learning, deep convolutional neural network architecture

INTRODUCTION

Pathogenic bacteria, physical and chemical causes, immunological damage, and numerous drugs cause inflammation of the lung parenchyma. Pneumonia classification encompasses several distinct approaches: (1) it is primarily categorized into infectious and non-infectious types based on pathogenesis, with infectious pneumonia being more prevalent. This group comprises radiation pneumonia, immune-associated pneumonia, and aspiration pneumonia due to chemical and physical causes. Non-infectious pneumonia encompasses conditions caused by immune responses, physical and chemical aspiration, and radiation exposure. (2) Community-acquired pneumonia (CAP), hospital-acquired pneumonia (HAP), and ventilator-associated pneumonia (VAP) are the three primary forms of the disease, with CAP accounting for the majority of cases. Owing to the diverse etiologies involved.

***Author for Correspondence**
Saniya P.M.
E-mail: saniya_cs@pace.edu.in

[1]Assistant Professor, Department of Computer Science and Engineering, PA College of Engineering, Mangalore, Karnataka, India
[2]Student, Department of Computer Science and Engineering, P A College of Engineering, Mangalore, Karnataka, India

Received Date: August 10, 2024
Accepted Date: August 26, 2024
Published Date: September 06, 2024

Citation: Saniya P.M., Mohammad Shabeer, Muhammed Sinan M.S., Muzammil Rahman K.M., Sanad Fazal Kota. Pneumonia Detection and Classification Using Deep Learning. Journal of Microwave Engineering & Technologies. 2024; 11(3): 9–19p.

HAP makes it simpler for bacteria to develop antibiotic resistance, which makes treatment more difficult. Nearly 800,000 children under the age of five die of pneumonia each year. More than 2200 individuals were killed. Pneumonia affects more than 1400 children out of 100,000 children. The Global Burden of Disease Study indicated that pneumonia and other respiratory disorders were the second most common cause of death in 2013. The incidence of pneumococcal disease is 35% in European hospitals and 27.3% globally. In 2015,

© STM Journals 2024. All Rights Reserved

India reported the highest mortality rate from pneumonia among children under five years of age, with nearly 297,000 combined deaths from pneumonia and diarrhea. Pneumonia mortality rates are inversely correlated with age, increasing notably in individuals over 65 years of age. The significant impact on infant mortality has spurred global efforts to enhance diagnostic and treatment methods, particularly through radiology-based approaches, such as chest X-ray imaging. Despite its advantages in accessibility and radiation dosage, diagnosing pneumonia from X-ray images remains challenging owing to overlapping features with other lung conditions, such as lung cancer. This study proposed a convolutional neural network (CNN) for autonomous pneumonia diagnosis from X-ray images, achieving an accuracy of 96.07% and an AUC of 0.9911. The subsequent sections outline the literature review on medical image processing, CNN architecture, and training methodology, followed by the experimental results and conclusions. The second section discusses the ideas found in the literature on medical photo-processing processes. A quick overview of the CNN architecture is provided in Section 3. The final section provides a summary of the evolution of deep learning and machine learning. Section 4 details the methodologies and techniques utilized in this study, including the materials employed and the proposed approach. Section 5 presents the experimental trials and their respective outcomes. The researcher's opinions and viewpoints on the study's conclusions and ramifications are summarized in Section 6.

RELATED WORKS

There are various challenges in developing a long-term version of COVID-19 detection using X-ray images. First, only a limited number of photos are available. As the epidemic has only been spreading for a few months, insufficient datasets have been collected and released for scientists to use. Second, there is a pressing need for scientists to develop an intelligent device capable of rapidly identifying and analyzing viral infections using chest X-ray images, which can provide crucial insights into the severity of the condition in the human body. After evaluating the current technologies and the challenges at hand, it was determined that transfer learning is a viable and practical approach to addressing this research problem. This approach involves using datasets of pneumonia images to train deep neural networks to detect illnesses affecting one or both lungs that can be caused by bacteria, viruses, or fungi. These conditions lead to inflammation in the alveoli and small air sacs in the lungs, impairing breathing as the respiratory system fills with fluids or pus. To identify the type of virus, the approach employs transfer learning, leveraging established deep learning methodologies that have been extensively studied and validated on diverse datasets. This process involves transferring the models and insights gained from the training and validation phases to a new dataset, specifically targeting a rapidly spreading infectious disease caused by coronaviruses, such as COVID-19, caused by the coronavirus. This method addresses the challenge of developing robust models from limited and relatively unexplored datasets, aiming to efficiently uncover and understand the unique properties of the virus [1].

Deep CNNs tend to achieve better performance with larger datasets than with smaller ones. Despite the widespread prevalence of COVID-19 globally, publicly available chest X-ray images are scarce, specifically for COVID-19. Therefore, to address this limitation, the authors of this study created a substantial dataset comprising COVID-19 chest X-ray images. This dataset augmentation was necessary given the abundance of available normal and pneumonia images, which were also utilized in this research [2].

The study utilizes various advanced deep learning techniques, employing six CNN models: GoogLeNet, LeNet, VGG-16, AlexNet, StridedNet, and ResNet-50. These models were trained on a dataset of 28,000 chest X-ray images, each resized to 224 × 224 pixels. The training is conducted with batch sizes of 32 and 64, aiming to assess and compare the accuracy of each model in predicting and distinguishing between affected and unaffected patients based on the chest X-ray images Adam is also used in the study as an optimizer, providing a 500-epoch learning rate and an adjusted 1e-four learning rate to each model. A 98 percent accuracy rate was obtained for both GoogLeNet and LeNet models during development, a 97 percent accuracy rate was obtained for VGGNet-16, an average accuracy rate of 96%%was obtained for the AlexNet and StridedNet models, an 80% accuracy rate was obtained for

© STM Journals 2024. All Rights Reserved

the ResNet-50 model comprehensive performance training, and GoogleNet and LeNet had the highest average accuracy [3]. The six models that were identified were able to recognize and anticipate the symptoms of pneumonia, including a normal chest radiograph.

The authors of this paper categorize X-ray images into three types: COVID-19, pneumonia, or normal. They utilized an ensemble technique during the prediction process, in which each image underwent classification through the respective layers dedicated to identifying these three categories. This approach ensures that each X-ray image is accurately classified based on whether it shows signs of COVID-19 infection or pneumonia or appears normal [4].

In the initial stage of statistical pre-processing, the X-ray dataset was partitioned into training and test subsets. During this phase, the chest X-ray images were normalized. Data augmentation techniques were implemented in the training images to improve the resilience of the model. The next step involves building a model to distinguish between consolidation and non-consolidation in the X-ray images. The performance of the model was evaluated by K-fold cross-validation, where k is the number of folds. Finally, an explainable AI technique was implemented to enhance the interpretability of the model. One way to achieve this is to create heatmaps, which are essentially visual aids for explaining how the model makes decisions. To assess the reliability of the model, two distinct methods were employed: 1) generating a heatmap using a single model and 2) producing heatmaps using an ensemble of models trained on different folds of the dataset. The ensemble approach allows for the calculation of uncertainty levels for each pixel, represented by standard deviations, thereby assessing the robustness and reliability of the heatmap [5].

In this study, the author proposes a modern approach utilizing an ensemble of Retina Net and Mask R-CNN to detect pneumonia regions in chest X-ray images. Pneumonia regions typically appear small, with mean dimensions of 304 pixels in height (29.6% of the image height) and a median width of 219 pixels (21.3% of the image width), as depicted in Figure 1. Traditional object detectors struggle to efficiently identify such small areas. To address this, the study employed a Feature Pyramid Network (FPN) as the backbone for both models. By offering multi-scale feature maps and combining enhanced semantics with low-resolution features via lateral connections and top-down routes, the FPN improves detection. Additionally, residual networks are utilized as the base architecture to mitigate the degradation issues associated with deeper networks, enabling the development of robust and effective models [6].

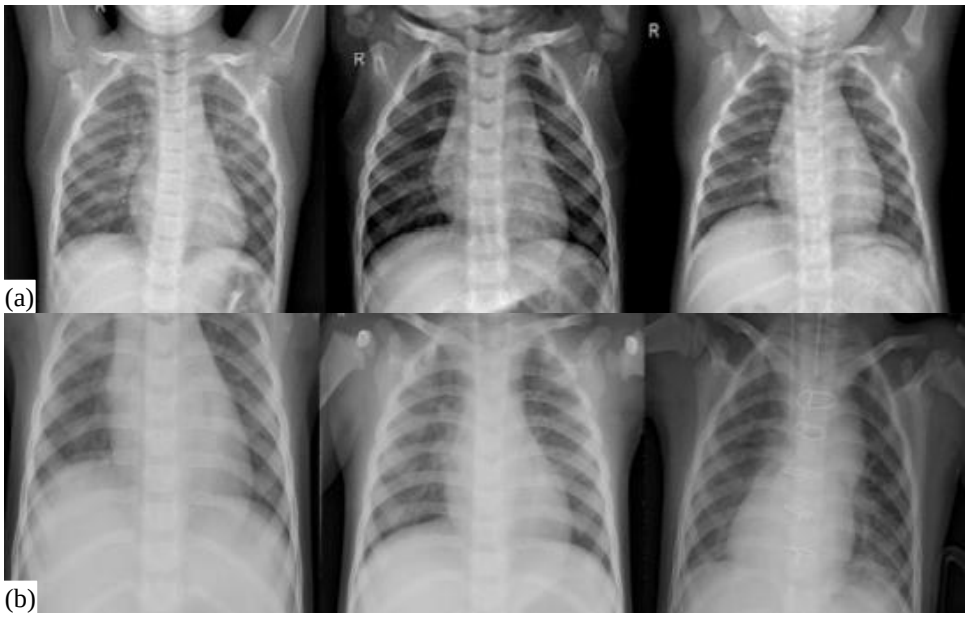

Figure 1. Examples from the dataset (a) normal cases, (b) pneumonia cases.

© STM Journals 2024. All Rights Reserved

BACKGROUND

In recent years, there has been growing interest in machine learning (ML) methodologies. This approach harnesses the significant potential of leveraging predetermined algorithmic phases for image processing. In contrast, traditional ML methods for partitioning tasks require manually designed algorithms or output layers to classify images. In addition, LeCun et al. [7] introduced a CNN approach capable of automatically extracting features by iteratively applying convolutional layers to the input images. These networks begin with shallow layers that focus on low-level image features. As the number of layers increases, CNNs capture increasingly complex features through backpropagation, adjusting the learned parameters to distinguish between different image categories. In CNNs, convolutional kernels filter the preceding image components, followed by pooling functions to reduce feature map dimensions and enhance computational efficiency. Nonlinear activation functions then process the resulting components to generate deeper insights into the image characteristics.

Integration tasks such as mid-level and high-level integration play crucial roles in enhancing the model's simulation capabilities. The integration process involves dividing the input into multiple regions, each varying in size both horizontally and vertically. High integration computes the average of each subregion, distinguishing it from mid-level integration. Activation functions, such as Rectified Linear Units (ReLU) and sigmoid, are commonly used to process image elements extracted through segmentation and convolutional operations, followed by integration and other fully integrated layers. This approach facilitates the extraction of pneumonia from processed images by evaluating derived features at the pixel level, thereby enhancing the overall capacity of the model. Over the past decades, significant neural frameworks, such as AlexNet [8] and VGGNet [9], have shaped deep learning advancements. However, increasing the network depth can lead to challenges, such as overfitting and gradient vanishing, limiting its capacity to learn productive features across diverse datasets. To address these issues, residual connections have been introduced to maintain an effective gradient flow and mitigate the degradation problem in deeper networks. This study explored the efficiency of residual connections in a simplified CNN architecture with fewer layers [10], demonstrating their utility in maintaining model performance while overcoming depth-related challenges.

MATERIALS AND METHODS

Data

The 5863 X-ray images in the suggested database, which will be used to evaluate the model's performance, were gathered using Kaggle. Dr. Paul Mooney organized a Kaggle competition in 2017 to differentiate between bacterial and viral pneumonia. It stands out from the other datasets because it includes 5,863 pediatric photographs. We discuss the most recent version of this dataset [6].

Three primary folders—train, test, and val—each holding a subfolder for distinct image categories (pneumonia and general) comprise the organization of the dataset. Figure 1 displays examples of both pneumonia and general chest X-ray images standardized to a fixed size. Because of the nature of chest X-ray imaging, which typically results in low-contrast images, the images exhibit varying shades of black, white, and gray. The lungs, located on either side of the thoracic cavity, appear dark on X-rays because of their air-filled structures. The heart, situated between the lungs, appears white as the X-rays penetrate completely. Bones, composed of dense material that X-rays cannot penetrate, appear very white on the images with clear and defined edges.

Data Pre-processing

Table 1 presents the strategies used in this study. Rescale is the number we multiply the data by before performing any additional analysis as part of our research. Because values such as this would be too large for our models to handle at the usual learning rate, we reduced the size of our original pictures by a ratio of 1. /255. The RGB coefficients in the original photos range from 0 to 255 [11]. Shearing transformations were randomly applied using the shear range. The zoom range was used to randomly zoom inside photos when there were no presumptions of horizontal asymmetry, and a horizontal flip

© STM Journals 2024. All Rights Reserved

Table 1. Data pre-processing techniques.

Rescale	1./255
Zoom Range	0.2
Shear Range	0.2
Horizontal_Flip	True

was used to randomly flip half of the images horizontally (e.g., real-world pictures). Data pre-processing techniques used in this study.

Proposed Network

In this study, we developed a CNN model based on VGG to extract features from chest X-ray images and utilize those characteristics to determine whether a patient has pneumonia. We began with a lower filter setting of 32 in our CNN architecture and worked our way up, layer by layer. The model was constructed using a MaxPooling layer and a Conv2D layer. An odd number, such as 3 × 3, is desirable for the kernel size [12].

The mechanisms of ReLU activation are the most widely used; however, Tanh and other methods can also be used. The width and height of an image are accepted by the input form, and the final dimension acts as a color channel. The input was then flattened, and artificial neural network (ANN) layers were added.

$$S(x)\,\frac{1}{1+e^{-x}} \tag{1}$$

f(x)=max(0, x)
S(x) = Sigmoid
f(x) = ReLU

For the last layers (ANN layers), I used a SoftMax activation function and defined the units as the total number of classes. Details of proposed DL model. Shown in Figure 2. For binary classification, a sigmoid was used, and the unit was set to 1.

IMPLEMENTATION

One type of neural network that possesses every trait of every other type is the CNNs. CNN, on the other hand, was designed with input image processing in mind. As a result, their organizational architecture is more focused on and consists of two main components.

Conv Layers

The initial block of a CNN plays a crucial role as a feature extractor, distinguishing it from traditional neural networks. This block operates by employing convolutional filtering operations similar to template matching. In this process, the first layer convolves the input image with multiple convolution kernels, generating what is known as "feature maps." These feature maps encapsulate the localized patterns and features detected in the input image. Subsequently, these maps undergo normalization through activation functions, such as ReLU or sigmoid, which enhance the nonlinear properties of the network and facilitate better gradient propagation during training. Additionally, feature maps may be downsampled through techniques such as max pooling to reduce spatial dimensions while retaining essential features.

This initial stage sets the foundation for subsequent layers to extract increasingly abstract and complex features as the information flows deeper into the network. CNNs capture hierarchical representations in images by systematically applying convolution and activation operations, making them effective for tasks such as image classification, object detection, and segmentation.

Pool Layers

The second block in a neural network, including CNNs, plays a critical role in the final stages of classification tasks. Unlike the unique feature extraction process of the initial CNN block, this stage is

© STM Journals 2024. All Rights Reserved

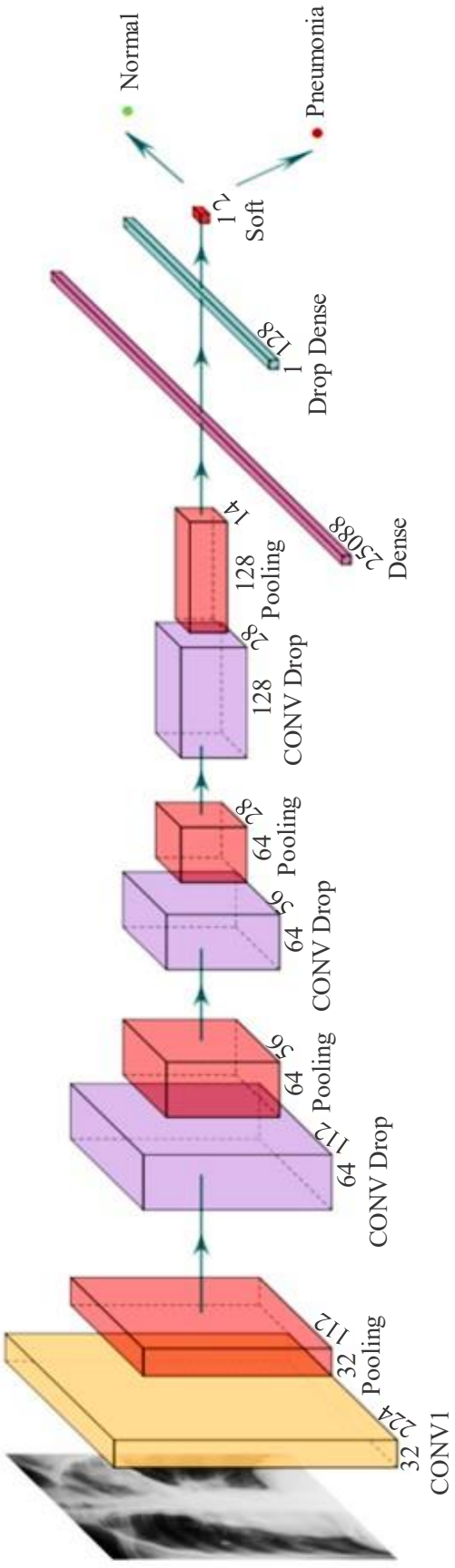

Figure 2. Details of proposed DL model.

© STM Journals 2024. All Rights Reserved

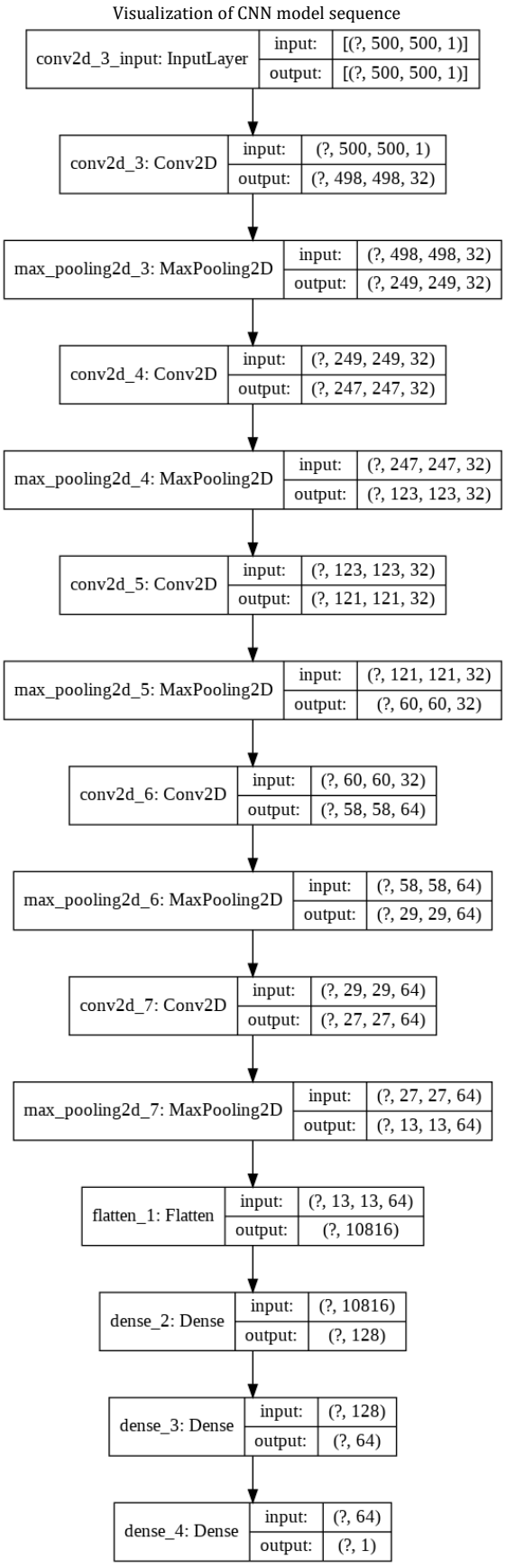

Figure 3. Visualization of CNN model sequence.

© STM Journals 2024. All Rights Reserved

common across various neural network architectures. The input vector undergoes transformations through linear combinations and activation functions to produce an output vector. In the context of classification, this output vector contains the probabilities for each class, indicating the likelihood that the input image belongs to each respective class.

For binary classification problems, a sigmoid activation function is typically employed in the final layer. The probability that a picture belongs to the positive class is represented by each value in the probability range of zero to one that this function outputs. In contrast, for multiclass classification, the SoftMax activation function was used. It normalizes the outputs across all classes, ensuring that they sum to 1 and thereby represent valid probability distributions.

This final layer's calculations are crucial as they determine the network's predictive accuracy, enabling it to assign class labels based on the highest probability score among the elements of the output vector. Visualization of the CNN model sequence is shown in Figure 3.

Fitting Model

EarlyStopping is a mechanism used during model training to prevent overfitting by halting the epochs based on a specified metric. When the validation loss in this instance did not continue to decline, training was terminated. The "patience" parameter determines how many epochs the training can continue after the validation loss stops decreasing before it stops.

In the training session, early stopping stopped the training process in the 9th epoch. At this point, the validation loss was 39.8%, and the validation accuracy was 68.7%. This means that after achieving a minimum validation loss if the loss increased in the subsequent three epochs (as per the patience level of 3), training would be halted at the 9th epoch to prevent further overfitting of the model. The model's actual test results were 91.98 percent accurate.

Visualizing some predicted images with percentage (%): Figures 4–12 reveal the prognosis of pneumonia using ML algorithms. This provides an estimate of the specific image as a percentage. The image can be loaded directly from the hard drive by providing its path.

- After importing the image as we previously did, the test set was entered into the model, and a forecast was created by repeating all the data pretreatment procedures. Adding the tensor flow too. Pre-processing requires an image class that is necessary for karas.
- Import a grayscale color channel image with dimensions of (500,500).
- To forecast the scenario, as previously demonstrated, the image was transformed into an array, rescaled by dividing it by 255 and extending the dimension by axis = 0.

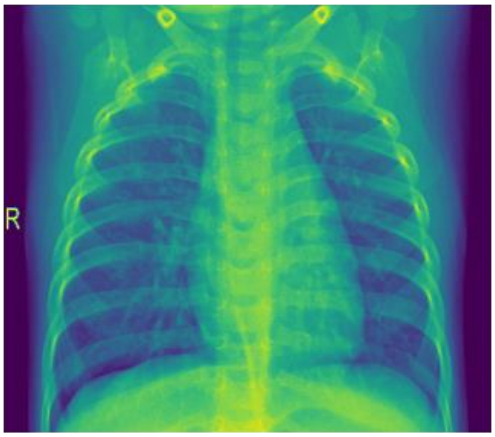

Figure 4. 95.54% probability of being normal case. Actual case: normal.

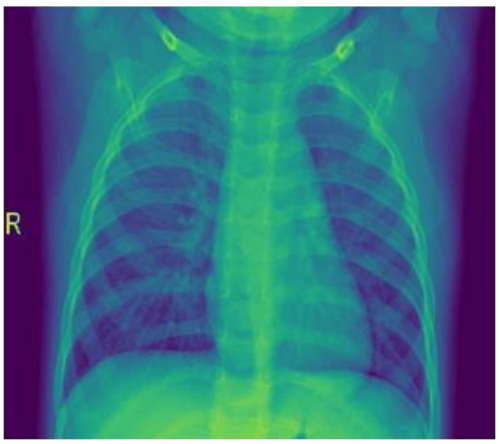

Figure 5. 97.97% probability of being a normal case. Actual case: normal.

© STM Journals 2024. All Rights Reserved

Figure 6. 81.72% probability of pneumonia case. Actual case: pneumonia.

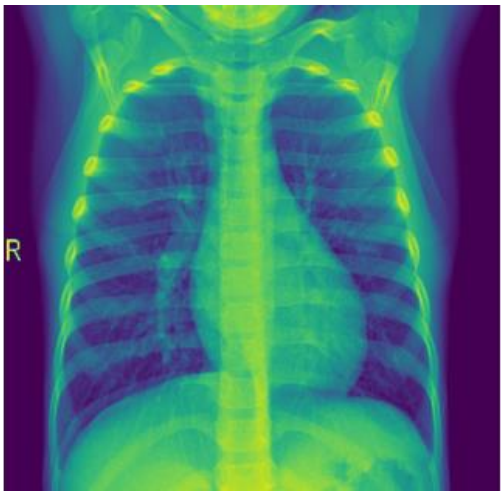

Figure 7. 99.27% probability of being normal case. Actual case: normal.

Figure 8. 99.05% probability of being normal case. Actual case: normal.

Figure 9. 98.74% probability of being a pneumonia case. Actual case: pneumonia.

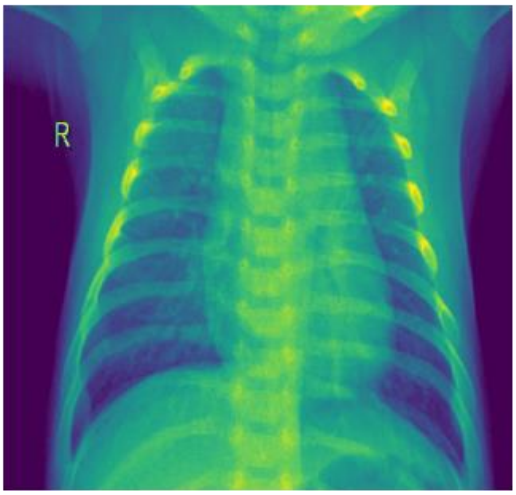

Figure 10. 91.18% probability of pneumonia case. Actual case: normal.

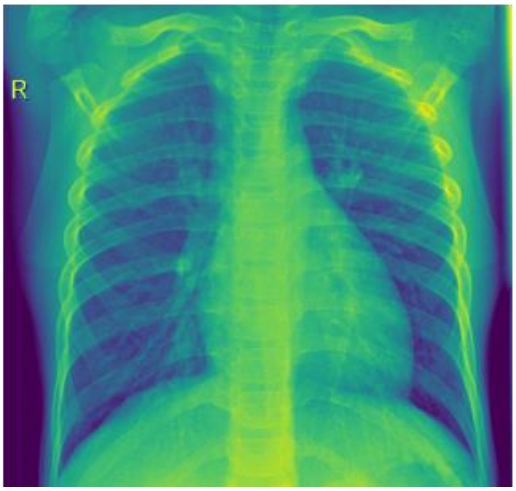

Figure 11. 90.59% probability of being a normal case. Actual case: normal.

© STM Journals 2024. All Rights Reserved

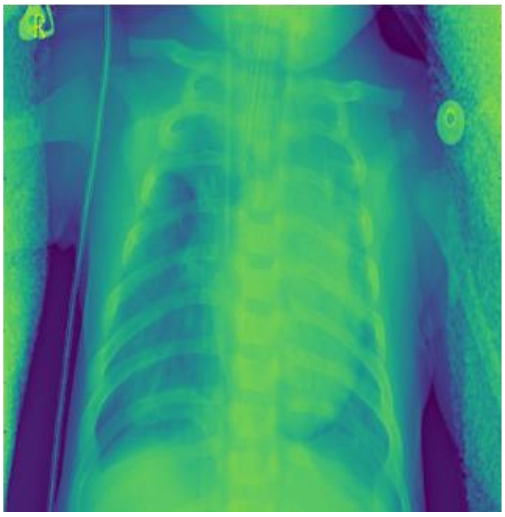

Figure 12. 99.99% probability of pneumonia case. Actual case: pneumonia.

CONCLUSION AND FUTURE WORK

This study introduces an automated approach for diagnosing pneumonia from X-ray scans using deep CNNs. The study leveraged the X-ray scan dataset comprising 5863 scans, achieving notable metrics, such as accuracy, recall, precision, and AUC ranking. The proposed framework demonstrated a high categorization accuracy of 91%. To enhance the model performance, hyper-parameter optimizations were explored, including stochastic gradient descent, Adagrad, and the Adam optimizer. The customized VGG16 model proved particularly effective in pneumonia detection, outperforming other optimizers, such as Adam. Future research aims to extend this work to detect and classify multiclass X-ray images and incorporate advanced feature extraction techniques from recent biomedical image segmentation models to further improve efficiency.

REFERENCES

1. Fernandes V, Junior GB, de Paiva AC, Silva AC, Gattass M. Bayesian convolutional neural network estimation for pediatric pneumonia detection and diagnosis. Comput Methods Programs Biomed. 2021;208:106259. DOI: 10.1016/j.cmpb.2021.106259. PubMed: 34273674.
2. Lu H, Hewakankanamge SA, Miao Y. Transfer learning from pneumonia to COVID-19. In: IEEE; 2020. pp. 1–6. DOI: 10.1109/CSDE50874.2020.9411550.
3. Militante SV, et al. Pneumonia and COVID-19 Detection using Convolutional Neural Networks. 2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE), Surabaya, Indonesia, 2020, pp. 1–6. DOI: 10.1109/ICVEE50212.2020.9243290.
4. Militante SV, Dionisio NV, Sibbaluca BG. Pneumonia detection through adaptive deep learning models of convolutional neural networks. 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 2020, pp. 88-93, doi: 10.1109/ICSGRC49013.2020.9232613.
5. Hasan MhdJ, Alom MS, Ali MS. Deep learning-based detection and segmentation of COVID-19 & pneumonia on chest X-ray images. 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), Dhaka, Bangladesh, 2021, pp. 210-214. DOI: 10.1109/ICICT4SD50815.2021.9396878.
6. Mooney P. (2018). Chest X-Ray Images (Pneumonia). Kaggle.com. [online] DOI: https://doi.org/10/23340/c8372ebbe20b0f671c2f3c501ba51412. Available from: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia
7. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten ZIP code recognition. Neural Comput. 1989;1(4):541–51. DOI: 10.1162/neco.1989.1.4.541.

© STM Journals 2024. All Rights Reserved

8. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90. DOI: 10.1145/3065386.
9. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. [Preprint]. arXiv 2014. Available from: arXiv:1409.1556.
10. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy. 2017. pp. 618–26. DOI: 10.1109/ICCV.2017.74.
11. Wang L, Lin ZQ, Wong A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci Rep. 2020;10:19549. DOI: 10.1038/s41598-020-76550-z.
12. Deshmukh H. (2020). Medical X-ray. Image Classification using Convolutional Neural Network. [online] Medium. Available from: https://towardsdatascience.com/medical-x-ray-%EF%B8%8F-image-classification-using-convolutional-neural-network-9a6d33b1c2a.

© STM Journals 2024. All Rights Reserved