Study and Analysis of Different CNN Architectures by Detecting Covid-19 and Pneumonia from Chest X-Ray Images

Manasa K M [1], Jesmitha Thoras[2]

[1,2] *Student, Computer Science and Engineering, Vivekananda College of Engineering and Technology, Puttur, Karnataka, India*

Abstract - *Covid-19, stemming from the Coronavirus, emerged in Wuhan, China, in 2019. Its rapid global spread led to approximately 5 million deaths. Pneumonia, an infection inflaming lung air sacs, affects one or both lungs. Diagnostic tools include blood tests and pulse oximetry for Pneumonia, and Swab tests, Nasal aspirates, Tracheal aspirates, and Sputum tests for Covid-19. However, these tests, while accurate, pose challenges in implementation and social distancing, especially in underdeveloped nations.*

Deep learning models have revolutionized healthcare with their computational prowess. These networks are reshaping patient care and crucial to modern clinical practices. This study assesses eight CNN architectures, determining ResNet-50 as the optimal choice. The dataset comprises 9208 chest X-Ray images across three classes: Covid-19, Normal, and Pneumonia. These images are divided with 60% for training, 30% for validation, and 10% for testing. The system receives input images, classifies them, and presents results with associated probabilities.

Key Words: **Pneumonia, Covid-19, Deep Learning, Convolution Neural Network, ResNet-50/ResNet-101, VGG-16/VGG-19, MobileNet.**

1. INTRODUCTION

COVID-19 is caused by SARS-CoV-2, an acute respiratory syndrome coronavirus. It originated in December 2019 in China and has since escalated into a pandemic. According to reports, there have been over 130 million cases worldwide [1].

The coronavirus, known as COVID-19, was declared a global health crisis and pandemic by the World Health Organization due to the extensive global impact it has had and its high level of contagion [2]. The coronavirus can be contracted through breathing droplets released by a person who is talking, sneezing, or coughing [3]. It spreads rapidly through close interaction with infected persons or by contact with contaminated surfaces and objects [4]. The best way to protect oneself from the virus is by avoiding exposure [4]. Pneumonia can be a life-threatening illness if not properly diagnosed and can lead to death in those afflicted with this ailment [5]. It presents as a severe respiratory illness caused by transmissible agents like viruses or bacteria that affect the lungs [5]. It can be spread through the nose or throat and can affect the lungs when inhaled or transmitted through airborne droplets from a person coughing or sneezing [6]. The lungs of a person are composed of small sacs called alveoli that facilitate the passage of air whenever a healthy individual breathes [5].

When a person is infected with pneumonia, it limits the oxygen intake and makes breathing difficult and painful due to tissue soreness caused by alveoli covered with fluids or pus [6]

While there are physical tests available to detect these diseases, it is difficult for some countries to conduct tests and maintain social distance during the pandemic. There are lot of research done on COVID-19 detection using multimodal images including CT, X-rays [7][8]. Some research even claims that detecting COVID-19 from multimodal images using artificial intelligence have more sensitivity as compared to RT-PCR test [9].

Machine learning and deep learning methods have become revolutionary from past few years in the field of computer vision and medical image processing for both classification and prediction problems [10]. The scarcity of samples available for positive cases makes this area of research more challenging. Concepts of transfer learning and pre-trained models have been used in previous researches which showed promising results, although these methods have not been yet used as alternative methods as they were experimentally effective but practically ineffective. [11][12].

2. Related Work

For the past few decades, computer-aided diagnostics (CAD) for medical imaging have been in use. A revolution is evident, transitioning from conventional methods to artificial intelligence approaches, and from machine learning to deep learning methods. Various techniques have been employed to classify diseases such as cancer, skin disorders, chest ailments, and many more.

Sammy V. [13] introduced the detection of Pneumonia and COVID-19 through the utilization of Convolutional Neural Networks. The VGG-16 architecture was employed for this purpose. The study aimed to identify the conditions by applying image processing techniques. The proposed CNN architecture includes nodes with weights that are updated during model training using optimization techniques like backpropagation. The study utilized a learning rate of 0.0001 with the Adam optimizer and cross-entropy using categorical features for optimization. The model achieved an accuracy of 95% in the study.

S. V. Militante and B. G. Sibbaluca conducted a study in which five distinct deep learning models were trained and compared. The study aimed to identify the best model for detecting pneumonia and healthy chest X-ray images. The results of the study revealed an accuracy rate of 97%, with the VGG-Net model being identified as the most effective in detecting pneumonia [14].

D. P. Yadav [15] introduced a Deep Learning Approach for the Detection and Classification of Bone Fractures. The experiment utilized X-ray images of both fractured and healthy human bones. To address the issue of overfitting, a data augmentation technique was employed to enhance the size of the dataset. The proposed model achieved a classification accuracy of 92.44% for distinguishing between healthy and fractured bones. Further enhancement in model accuracy could be achieved by considering alternative deep learning models.

Shelly Soffer [16] introduced Convolutional Neural Networks for Radiologic Images. The research outlines the stages involved in designing deep learning approaches for radiology. The article elaborates on a survey that explores the application of deep learning, particularly Convolutional Neural Networks, in radiologic imaging, with a focus on five major organ systems: chest, breast, brain, musculoskeletal system, abdomen, and pelvis. The utilized CNN architectures include AlexNet, VGG16 and U-Net, which assist in classification. The proposed model achieved a classification accuracy of 85%.

3. METHODOLOGY

This section provides an overarching description of the applied methodology. In this study, various architectures available in Convolutional Neural Networks (CNN) were utilized. The proposed diagnostic approach for Covid-19 and Pneumonia involved the following steps:

1. Dataset Collection and Description
2. Dataset Splitting
3. Data Pre-processing
4. Feature Extraction/Model building
5. Model Training

1. Dataset Collection and Description

The dataset was obtained from Kaggle [17]. It comprises a total of 9208 Chest X-Ray images categorized into three classes: Normal, Covid-19, and Pneumonia. Among the entire dataset, there are 3270 images depicting Normal Chest X-rays, 1281 images showcasing Covid-19 affected Chest X-rays, and 4657 images representing Pneumonia Chest X-rays. All images possess dimensions of 299x299 pixels. Sample images from the dataset are illustrated in Fig.1 below.

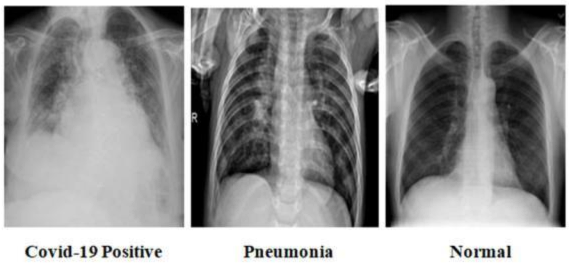

Fig – 1: Sample Chest X-Ray Images

2. Dataset Splitting

In this research, 9208 images from the dataset were utilized to develop and assess the model. The dataset was divided in a ratio of 6:3:1, where 60% of the dataset was allocated for training, 30% for validation, and the remaining 10% for testing.

3. Data Pre-processing

The primary objective behind employing Convolutional Neural Networks (CNN) in many image classification tasks is to reduce the computational complexity of the model. The pre-processing steps outlined in this methodology are as follows:

Initially, the original images were resized from 299x299 pixels to 224x224 pixels. This resizing was done to alleviate computational load and enhance processing speed. All subsequent techniques were applied to these resized images.

Data augmentation, a common practice in the medical domain, was employed to expand the dataset. Various data augmentation methods such as rotation, horizontal and vertical flipping, and zooming were utilized. These techniques aid in extracting more features from the images.

For the architectures including ResNet50, ResNet101, DenseNet, Xception, MobileNet, VGG-16, and VGG-19, the built-in Keras feature "preprocess_input" was utilized to align with the model's requirements. In the case of AlexNet, the rescale technique was employed.

4. Model Building – CNN

The Convolutional Neural Network (CNN) is built using a multitude of smaller units called nodes, organized within a layered architecture. These nodes are comprised of weights that are updated during the model's training using optimization techniques like backpropagation [18]. CNN excels in image classification tasks and delivers high accuracy.

Every CNN architecture predominantly comprises various layers, including the convolution layer, pooling layer, batch normalization, fully connected layer, dropout layer, and activation function. Figure 2 illustrates the distinct layers of the CNN.

a. Convolution Layer: Serving as the initial layer, this component is utilized to extract diverse features. Through a mathematical operation known as convolution, a filter of specific size MxM is slid over the input image. This entails taking the dot product between the filter and portions of the input image aligned with the dimensions of the filter (MxM) [19]. The result of this operation is a feature map, containing information about image patterns. Subsequently, this map is passed as input to the subsequent layer.

b. Pooling Layer: The main purpose of this layer is to reduce the size of the feature map, thereby reducing computational costs [19]. Several types of pooling techniques exist, such as Max Pooling and Average Pooling. For our study, we predominantly utilized Max Pooling, where the largest element is selected from the feature map. Additionally, we employed Average Pooling, which involves calculating the average of the elements within the feature map.

c. Batch Normalization: This layer is employed to enhance the learning rate of the CNN model and standardize the input image. Within a CNN model, batch normalization is implemented after each convolutional layer [18].

d. Dropout Layer: The primary purpose of this layer is to mitigate the issue of overfitting. During the training process, the dropout layer systematically omits a portion of neurons from the model, leading to a reduction in the model's size.

e. Fully Connected Layer: This layer accepts the outputs from preceding layers, flattens them, and transforms them into a single vector. These layers are positioned prior to the classification output of the model.

f. Activation Functions: These functions are employed to learn and approximate intricate relationships between variables within the network. Numerous activation functions exist, and for this study, ReLU and SoftMax functions were utilized. ReLU introduces non-linearity to the network and serves as a solution to the vanishing gradient problem. SoftMax is applied for multiclass classification tasks.

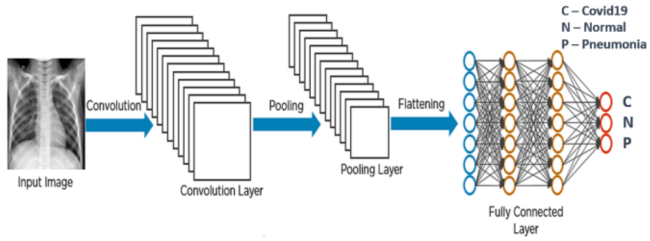

Fig – 2: CNN Layers

The analysed architectures are,

i. AlexNet: AlexNet is a deep convolutional neural network (CNN) that made significant advancements in the field of computer vision and image recognition. It was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, and it won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012, marking a breakthrough in deep learning. AlexNet consists of five convolutional layers followed by three fully connected layers. It employs rectified linear units (ReLU) as activation functions, which helped alleviate the vanishing gradient problem. AlexNet's architecture and success in the ILSVRC competition marked a turning point in the field of deep learning, catalyzing the development of more complex and accurate neural network models for a wide range of applications beyond image recognition.

ii. DenseNet121: DenseNet121 was collaboratively developed by Cornell University, Tsinghua University, and Facebook AI Research (FAIR). A DenseNet is a kind of convolutional neural network that employs dense connections between layers using Dense Blocks. In these blocks, all layers with corresponding feature-map sizes are directly interconnected. DenseNet-121 comprises 1 layer with a 7x7 convolution, 58 layers with 3x3 convolutions, 61 layers with 1x1 convolutions, 4 Average Pooling layers, and 1 Fully Connected Layer.

iii. MobileNet: MobileNet, created by Michael Bartholomew, encompasses 27 Convolutional layers, which is a part of the efficient convolutional neural network (CNN) family designed for mobile and embedded devices with limited computational resources. These networks are specifically engineered to perform well on tasks such as image classification, object detection, and semantic segmentation while maintaining a small memory footprint and low computational requirements. MobileNet models are characterized by their ability to achieve a good balance between model accuracy and efficiency. This architecture comprises 13 depthwise Convolution layers, 1 Average Pooling layer, 1 Fully Connected layer, and 1 Softmax layer, making it suitable for a variety of real-time and on-device applications where computational resources are constrained.

iv. ResNet50/ResNet101: ResNet, introduced by He et al. [20], emerged victorious in the 2015 ImageNet competition. This methodology demonstrated the feasibility of training deeper networks. Interestingly, the accuracy of the network becomes more saturated as it grows deeper. This phenomenon is not solely attributed to overfitting or the abundance of parameters, but rather due to a decline in training error. This is a result of challenges in backpropagating gradients. This obstacle is surmounted by transmitting gradients directly to deeper layers using a residual block.

v. VGG16/VGG19: VGG16 and VGG19 correspond to architectures with 16 and 19 layers, respectively. These architectures were created by Karen Simonyan and Andrew Zisserman, securing the top position in the 2014 ImageNet challenge [21]. They are composed primarily of 3x3 convolutional layers stacked on top of each other, with occasional max-pooling layers for down-sampling. VGG16 and VGG19 are important milestones in the development of deep learning models for computer vision. Their simplicity, depth, and strong performance in the ImageNet challenge contributed to the advancement of deep convolutional neural networks and continue to influence the design of modern architectures.

vi. Xception : Xception, short for "Extreme Inception," is a highly advanced convolutional neural network architecture that consists of an impressive 71 layers, which was developed by François Chollet. The key innovation of Xception lies in its novel use of depthwise separable convolutions. This architectural choice effectively reduces the computational complexity of the network while preserving its ability to capture complex features from input data. The depthwise separable convolutions, inspired by MobileNet, break down the traditional convolution operation into two separate steps: depthwise convolution (applying a filter to individual input channels) and pointwise convolution (combining the outputs of the depthwise convolution with 1x1 convolutions). This approach dramatically reduces the number of parameters and computations compared to traditional convolutional layers, making Xception highly efficient.

5. Model Training

Upon constructing the model, the training phase was initiated. The optimizer used was Adam, chosen for its ease of implementation, computational efficiency, and low memory requirements. A batch size of 64 was specified, and the model was trained for 15 epochs with 35 steps per epoch. The early stopping technique was implemented to cease training when the model's performance plateaued on the validation dataset.

The convolution layers are equipped with trainable weights. Consequently, pre-trained weights from the ImageNet dataset were utilized, as they had undergone prior training.

4. EXPERIMENTAL RESULTS

Different architectures were assessed using the methodology outlined in section III. The performance of the evaluated deep CNN models was gauged through classification accuracy.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives of the classified Chest X-Ray images.

The X-ray image classification results for the used Dataset [17] for all evaluated deep CNN models after training them are tabulated in Table I. The best performance is indicated in bold.

Table – 1: Classification Accuracy (In percentages) for different trained Deep CNN Models in Multiclass COVID-19 and PNEUMONIA classification for the dataset.

	Accuracy		
Model Name	Training	Validation	Testing
AlexNet	82.28	70.50	74.35
DenseNet-121	94.60	92.38	93.01
MobileNet	94.82	92.96	93.40
ResNet-50	**95.85**	**94.14**	**94.76**
ResNet-101	95.18	93.65	94.48
VGG-16	94.06	92.96	93.23
VGG-19	93.66	89.16	91.36
Xception	91.16	86.91	89.65

Observing the accuracy table reveals that the top-performing model in terms of accuracy is ResNet-50, boasting a training accuracy of 95.85%, validation accuracy of 94.14%, and testing accuracy of 94.76%.

Results

Covid

Rank	Class	Probability
1st	Covid	86.32 %
2nd	Pneumonia	8.27 %
3rd	Normal	5.41 %

Results

Pneumonia

Rank	Class	Probability
1st	Pneumonia	95.24 %
2nd	Covid	4.7 %
3rd	Normal	0.06 %

Fig – 3: An example output of our system for Covid affected CXR (Left) Pneumonia infected CXR (Right)

5. CONCLUSIONS

The developed Convolutional Neural Network (CNN) models have proven effective in extracting meaningful features from X-ray images, facilitating the detection of COVID-19 and Pneumonia infections. Through the application of data augmentation techniques, feature extraction was bolstered across the training, validation, and testing phases.

As evident from the progress in computer-related applications within the medical domain, the integration of CNN and deep learning technologies has enabled the efficient detection of COVID-19 and pneumonia using chest radiographs. Among the various architectures evaluated, ResNet-50 emerged as the optimal choice for Covid-19 and Pneumonia detection, achieving an accuracy of 94.76%.

In the future, enhanced performance can be achieved through hyperparameter tuning, the exploration of diverse transfer learning combinations, and the utilization of Machine Learning algorithms for classification subsequent to feature extraction by CNN models. Furthermore, this model holds potential for identifying various other diseases such as Brain tumors and Bone fractures by training the model on the respective datasets.

REFERENCES

[1] https://www.worldometers.info/coronavirus

[2] World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report– 196,2020, https://www.who.int/docs/default-source/coronaviruse/situationreports/20200 803-covid-19-sitrep-196cleared.pdf?sfvrsn=8a8a3ca4_4

[3] Centers for Disease Control and Prevention. Interim Infection Prevention and Control Recommendations for Patients with Suspected or Confirmed Coronavirus Disease 2019 (COVID-19) in Healthcare Settings.2020. https://www.cdc.gov/coronavirus/2019-ncov/hcp/infectioncontrolrecommendations.html

[4] S. Militante, N. Dionisio, "Real-Time Facemask Recognition with Alarm System using Deep Learning".

[5] S. Militante, N.Dionisio, B.Sibbaluca, "Pneumonia Detection through Adaptive Deep Learning Models of Convolution Neural Networks".

[6] S. V. Militante, and B. G. Sibbaluca, April 2020,"PneumoniaDetectionUsing Convolutional Neural Networks", International Journal of Scientific & Technology Research, Volume 9, Issue 04, pp. 1332-1337.

[7] M.-Y. Ng et al., 2020"Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review," Radiology: Cardiothoracic Imaging, vol. 2, no. 1, p. e200034.

[8] H. Liu, F. Liu, J. Li, T. Zhang, D. Wang, and W. Lan, 2020/03/21/ 2020, "Clinical and CT imaging features of the COVID-19 pneumonia: Focus on pregnant women and children," Journal of Infection.

[9] Tao Ai, Zhenlu Yang, Hongyan Hou, Chenao Zhan, Chong Chen,Wenzhi Lv, Qian Tao, Ziyong Sun, and Liming Xia. 2020. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology (2020), 200642. https://doi.org/10.1148/radiol.2020200642 PMID:32101510

[10] Greenspan H, Van Ginneken B, Summers RM (2016) Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans Med Imaging 35:1153–1159.

[11] Shervin Minaee, Rahele Kafieh, Milan Sonka , Shakib Yazdani, Ghazaleh Jamalipour Soufi, October 2020, " Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning", ELSEVIER, Medical Image Analysis Volume 65.

[12] Ioannis D. Apostolopoulos, Tzani A. Mpesiana, 2020, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks", Physical and Engineering Sciences in Medicine 43:635–640 https://doi.org/10.1007/s13246-020-00865-4.

[13] Sammy V. Militante,Brandon G. Sibbaluca and Nanette V. Dionisio, 2020 ,"Pneumonia and COVID-19

Detection using Convolutional Neural Networks", 3rd International Conference on Vocational Educationand Electric Engineering (ICVEE).

[14] S. V. Militante, and B. G. Sibbaluca, April 2020, "Pneumonia Detection Using Convolutional Neural Networks", International Journal of Scientific & Technology Research, Volume 9, Issue 04, pp. 1332-1337.

[15] D. P. Yadav, Feb 2020. "Bone Fracture Detection and Classification using Deep Learning Approach" International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC) GLA University, Mathura, UP, India.

[16] Shelly Soffer, Avi Ben-Cohen, Orit Shimon, Michal Marianne Amitai, Hayit Greenspan and EyalKlang, 2019 "Convolutional Neural Networks for Radiologic Images", (RSNA).

[17] https://www.kaggle.com/datasets/francismon/curated-covid19-chest-xray-dataset

[18] Sammy V. Nanette V. Dionisio.2020 "Pneumonia and COVID-19 Detection using Convolutional Neural Networks" third International Conference on Vocational Education and Electrical Engineering (ICVEE).

[19] https://www.google.com/search?q=%5B19%5D+https%3A%2F%2Fwww.upgrad.com%2Fblog%2Fbasic-cnn-architecture%2F%23%3A~%3Atext%3DThere%2520are%2520thre+e%2520types%2520of%2CCNN%2520architecture%2520+will%2520be%2520formed.&oq=%5B19%5D%09https%3A%2F%2Fwww.upgrad.com%2Fblog%2Fbasic-cnn-architecture%2F%23%3A~%3Atext%3DThere%2520are%2520thre+e%2520types%2520of%2CCNN%2520architecture%2520+will%2520be%2520formed.&aqs=chrome..69i57.726j0j7&sourceid=chrome&ie=UTF-8

[20] https://arxiv.org/pdf/1512.03385.pdf

[21] Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide – 2019, Shelly Soffer, Avi Ben-Cohen, Orit Shimon,Michal Marianne Amitai,Hayit Greenspan, EyalKlang, M