International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

Classification of Chest X-ray Images using CNN for Medical Decision Support System

Pritam Garud[1], Gargi Patel[1], Nilkanth Kadlag[1], Sheetal Bhandari[1], Rachana Yogesh Patil[1]
Department of Electronics and Telecommunication Engineering
[1]Pimpri Chinchwad College of Engineering
Pune, India
pritam.garud45@gmail.com, gargipatel1501@gmail.com, nilkanthkadlag@gmail.com, sheetalubhandari@gmail.com,
rachana.patil@pccoepune.org

Abstract— X-rays are a crucial tool used by healthcare professionals to diagnose a range of medical conditions. However, it is important to keep in mind that a timely and accurate diagnosis is crucial for effective patient management and treatment. While chest X-rays can provide highly precise anatomical data, manual interpretation of the images can be time-consuming and prone to errors, which can lead to delays or incorrect diagnoses. To address these issues, healthcare systems have taken steps to improve diagnostic imaging services following the impact of the COVID-19 pandemic. While deep learning-based automated systems for classifying chest X-rays have shown promise, there are still several challenges that need to be addressed before they can be widely used in clinical settings, including the lack of comprehensive and high-quality datasets. To overcome these limitations, a real-time DICOM dataset, has been converted to JPEG format to increase processing speed and improve data control. Three pre-trained models and a convolutional neural network (CNN) model with low complexity and three convolutional layers for feature extraction, along with max pooling layers and ReLU and Softmax activation functions have been implemented. With an validation accuracy of 95.05% on their CNN model using the SGD optimizer, the result has been validated using a separate, real-time unlabeled DICOM dataset of 1000 X-ray images.

Keywords- Chest x-rays; Machine Learning; CNN Model; Deep Learning; InceptionV3; ResNet50.

I. INTRODUCTION

Medical image classification plays a crucial role in clinical treatment by aiding in the prediction of disease severity in patients, and supporting numerous research and diagnostic procedures. Although chest X-rays offer precise anatomical information, their manual interpretation can be prone to human errors and time-consuming, leading to delayed or inaccurate diagnoses. As the shortage of radiologists in the medical field has become a growing concern, especially with the increasing number of X-ray reports that need to be analyzed. This shortage can result in delays in diagnoses, as there may be a backlog of X-ray reports waiting to be read by a radiologist. In addition, these methods require a lot of time and effort for extracting and selecting classification features. Artificial Intelligence and Machine learning have facilitated the application of deep neural networks as effective automated classification tools in the field of medical diagnosis and convolutional neural networks consistently show the best results on varying image classification tasks. As medical images are typically delivered in DICOM (Digital Imaging and Communications in Medicine) format, which is a prevalent standard format that is utilized in the healthcare industry for both viewing and storing X-ray information. We understand how crucial it is to train and evaluate our CNN model using authentic data. Specifically, we are working with an original Chest X-ray dataset that is in DICOM format. There are many datasets available on internet today but using original data is crucial for maintaining the accuracy of machine learning models in medical imaging. by working with original Chest X-ray datasets and carefully converting them to JPEG format for use with our CNN model, we can achieve the highest level of accuracy in medical imaging. This approach ensures that our CNN model is well-equipped to handle real-world scenarios and provide accurate diagnoses that can lead to improved patient outcomes. To aid physicians and radiologists in the early detection and isolation of chest- related diseases in patients, the analysis has focused on accurately classifying chest X-ray images. For this purpose, a combination of a customized CNN model and three pre-trained models such as VGG16, Inception V3, and Resnet50 have been employed. This approach ensures that the classification results are reliable and can provide valuable insights to the medical professionals in the diagnosis and treatment of their patients. The customized CNN model was designed with a specific focus on classifying Normal and Abnormal Chest X- rays. By training this model on the transformed DICOM dataset, it is ensured that it is well-equipped to handle the specific nuances and complexities of this type of medical imaging. the comparison study of these models has been shown in table 1. In order to improve the accuracy of the classification of chest X-ray images, a variety of deep

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

learning models have been employed, including pre-trained models such as VGG16, Inception V3, and Resnet50. While these pre-trained models have been extensively used and have demonstrated high levels of accuracy in image classification tasks, the customized CNN model designed specifically for this project has outperformed them in terms of accuracy. Although the pre-trained models have also demonstrated high levels of accuracy in classifying chest X-ray images, the customized CNN model has proven to be the most effective for this particular task. Its accuracy and reliability make it a valuable tool for physicians and radiologists in the diagnosis and treatment of chest-related diseases in patients.

II. LITERATURE REVIEW

To accomplish this project, we extensively reviewed the following research papers. These papers offer a comprehensive understanding of the current advancements in the field and enable us to identify any existing gaps in the literature, as well as relevant concepts and approaches.

One of the papers titled "Chest X-Rays Image Classification in Medical Image Analysis" presents a detailed overview of data mining, medical image analysis, chest radiography, and machine learning. It aims to enhance readers' comprehension of the research domain. Notably, the utilization of machine learning techniques for disease detection and classification in chest X-ray images has demonstrated promising outcomes, with numerous studies being conducted in this domain. [1]

The paper titled "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks" focused on assessing the effectiveness of advanced convolutional neural network architectures in automatically detecting COVID-19 disease from X-ray images. Given the limitations of current diagnostic tests for COVID-19, utilizing X-ray images as a diagnostic tool may prove valuable, especially in situations where widespread or practical testing is not feasible. The study aimed to evaluate the performance of various convolutional neural network architectures and their potential for accurate detection of COVID-19. [2]

The literature survey titled "Medical image analysis based on deep learning approach" emphasizes the significance of utilizing deep learning techniques in the field of medical image analysis. It offers valuable information on recent advancements and potential applications of deep learning in various medical imaging modalities, such as X-ray images, computerized tomography, mammography images, and digital histopathology images. The survey sheds light on the growing importance of deep learning approaches and their impact on advancing medical image analysis. [3]

"Deep Learning in Multi-Class Lung Diseases' Classification on Chest X-ray Images" This paper proposes a deep learning method using the transfer learning technique to classify lung diseases on CXR images. The proposed method is a one-step, end-to-end learning approach that directly inputs raw CXR images into a deep learning model (Efficient Net v2-M) to extract meaningful features for identifying disease categories. However, further studies are needed to evaluate the generalizability of the proposed method and to validate its performance on larger and more diverse datasets. [4]

The paper titled "Deep Learning in Multi-Class Lung Diseases' Classification on Chest X-ray Images" introduces a deep learning approach that utilizes transfer learning to classify lung diseases based on chest X-ray (CXR) images. The proposed method employs an end-to-end learning process where raw CXR images are directly inputted into a deep learning model (Efficient Net v2-M) to extract relevant features for disease identification. However, additional research is necessary to assess the applicability of the proposed method across different datasets and to validate its performance on larger and more diverse samples. [5]

The article titled "Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis" examines the advancements made in object detection and semantic segmentation using convolutional neural network (CNN) algorithms in the field of medical imaging. The authors highlight that while medical imaging classification has been extensively researched, the exploration of object detection and semantic segmentation is relatively limited. The paper offers a comprehensive overview of CNN algorithms' application in object detection and semantic segmentation for medical imaging, emphasizing the precise identification of disease locations and boundaries. [6]

"Medical Image Classification and Cancer Detection using Deep Convolutional Neural Networks" This paper review lacks specific details on the methods used or results obtained. Nonetheless, it describes the importance of deep learning in medical imaging applications, particularly in medical image classification and cancer detection. It suggests that appropriate classifiers are needed for medical image classification and that transfer learning can be used for cancer detection. [7]

"Automated image classification of chest X-rays of COVID-19 using deep transfer learning" This paper proposes an early automated screening model for the detection of COVID-19 using chest X-ray (CXR) images. The authors use a database of 447 COVID-19 and 447 non-findings CXR images, which were divided into training, validation, testing, and local/Aligarh datasets. The authors used a Convolutional Neural Network

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

(CNN) architecture based on Mobile Net, which decreases computational cost while increasing speed. [8]

“Deep convolutional neural network-based medical image classification for disease diagnosis” This paper states that the traditional method of medical image classification has limitations in terms of performance and requires a lot of time and effort to extract and select classification features. CNN is a promising machine learning method for various classification tasks, especially in image classification, and the convolutional neural network dominates with the best results. However, medical image datasets are hard to collect due to the need for professional expertise to label them. [9]

The research paper titled "DICOM Imaging Router: An Open Deep Learning Framework for Classification of Body Parts from DICOM X-ray Scans" introduces an innovative tool designed to automatically classify body parts in DICOM X-ray scans using deep convolutional neural networks (CNNs). The authors acknowledge that the utilization of non-normalized DICOM databases in X-ray imaging poses a difficulty in implementing AI solutions for medical image analysis, as accurately identifying the relevant body part is crucial prior to inputting the image into a specific AI model. [10]

The aforementioned papers have investigated the utilization of machine learning and deep learning to manage and organize data, which results in enhanced diagnostic accuracy. After conducting a literature survey, it was discovered that CNN models yield higher accuracy in classifying X-rays. However, the previously mentioned papers had the limitation of solely using JPEG image datasets from unlabeled sources and applying different pre-trained models for image classification. Medical image classification is a significant process used to study, categorize, and predict health issues in humans. Although there are multiple classification models, the model that provides the highest accuracy is considered the best. The above papers solely relied on JPEG datasets from the internet to train and test the models, leading to a reduction in accuracy and loss of metadata. [11-12] To address this issue, this project aims to identify the most accurate model for DICOM to JPEG image conversion with minimal metadata loss. This will be achieved by employing real-time data to train the CNN model and classify chest X-rays, utilizing a deep learning approach.

III. PROPOSED METHODOLOGY

Overall structure of the conversion of dataset from DICOM to JPEG and classification based on abnormality prediction has been shown in Figure 1:

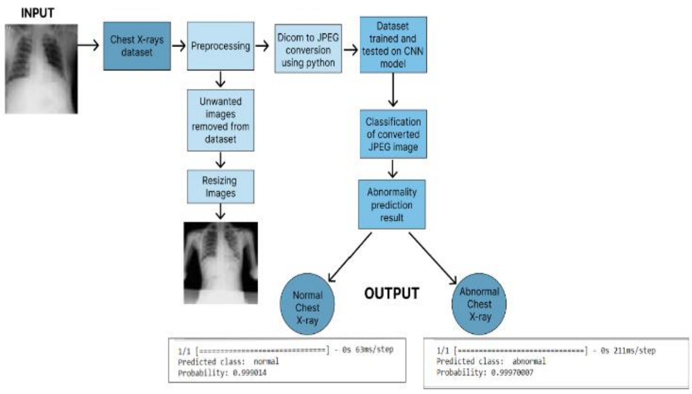

Figure 1. step-by-step classification and conversion process of CNN model.

The CNN model was trained using Chest X-ray images in the DICOM format, which is real-time data. The dataset comprised 2527 labeled DICOM images that were provided by Krsnaa Diagnostics Laboratory, Pune. Prior to training the CNN model, the DICOM format had to be converted into JPEG format. The metadata loss during this process was minimal due to the real-time nature of the data collection. The converted data was then subjected to various processes, including normalization, RELU function, Resblock, and average pooling, to extract image features. These processes enabled the identification of abnormalities in the chest X- ray image. The SoftMax block received the classes (Normal and Abnormal), which were used to categorize the image into either of the two classes. Therefore, this model is capable of classifying chest X-rays as either normal or abnormal.[13]

A. Convolution Neural Network Model

Convolutional neural networks (CNNs), a kind of artificial neural networks, are being employed more and more in radiology and other computer vision applications. CNNs use a variety of components, including convolution layers, pooling layers, and fully connected layers, to efficiently and automatically acquire hierarchical spatial features using backpropagation. The blockc diagram of CNN is shown in figure 2.

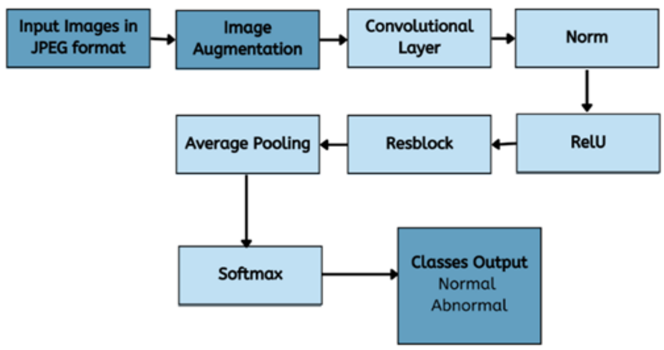

Figure 2. Block diagram of CNN model.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

- Input Images:

The collection consists of 2527 real-time DICOM Chest X-ray Images. This is a labeled dataset consisting of 1151 Normal X-ray images and 1376 abnormal X-rays images. The model receives these images as an input.

- Image Augmentation:

Computer vision and deep learning research rely on image augmentation. It enhances the training dataset by transforming the original images. Rescaling, shearing, zooming, rotation, flipping, and more are possible. Rescaling images normalizes pixel values, ensuring consistent input ranges and faster convergence during training. Shearing distorts the image along an axis, making the model orientation-invariant. Zooming images randomly simulates different perspectives and makes the model more robust to object size variations. Rotation lets the model learn from different orientations by rotating the image. Mirrored images are created by flipping images horizontally or vertically. When object orientation doesn't matter, these augmentations are helpful. Image augmentation can improve the model's generalization, overfitting resistance, and performance on unseen data. The model learns diverse and discriminative features for accurate classification and other computer vision tasks by adding variability to the training dataset. [14] The dataset and problem should determine the augmentation technique and parameters. Optimizing these methods often requires experimentation and fine-tuning.

- Convolution Layer:

Convolutional neural networks (CNNs) used in computer vision tasks depend on the convolutional layer. It helps detect local patterns and extract meaningful features from input data like images. Convolutional layers convolve input data with kernels, or filters. Each filter scans the input using a sliding window, performing element-wise multiplication and summation.[15] This allows the network to detect edges, textures, and more complex features at different spatial locations. CNNs learn hierarchical data representations using convolutional layers. Deeper layers detect patterns and structures from low-level features. Understanding visual data complexity requires hierarchical feature extraction.

- Norm:

Normalization standardizes data preprocessing. It involves standardizing diverse data sources. Neglecting to normalize data before training can slow our network's training and learning speed. A formula calculates the normalized value as a ratio of the data value difference and its range.

$$Xn = (X - Xminimum) / (Xmaximum - Xminimum)$$

Xn = Value of Normalization

Xmaximum = Maximum value of a feature

Xminimum = Minimum value of a feature.

- ReLU:

Convolutional layers used ReLU. ReLU, a popular activation function, introduces non-linearity into the network, allowing it to capture complex relationships between input data and features.[16] After each convolution operation, the ReLU activation function introduced non-linearity and enabled the model to acquire and represent more complex features. ReLU solves the vanishing gradient problem in deep neural networks. ReLU prevents the model's gradients from decreasing too much as they propagate backward, improving learning efficiency.

- ResBlock:

Each layer in a residual block adds its output to a layer below it. The output is combined with the main path output and passed through a non-linear activation function. This shortcut or skip connection bypasses network layers.

- Max Pooling:

Convolutional neural networks (CNNs) used in computer vision require the max-pooling layer. Its main function is to down sample input representation, reducing feature map spatial dimensions and benefiting the network. Max-pooling divides the convolutional layer output into non-overlapping regions, called pooling windows. The maximum value from each window is propagated forward. This process abstracts feature maps, capturing the most important features and discarding the rest.[17]

Max-pooling's translation invariance benefits the network. Selecting the maximum value within each pooling window makes feature location less important. The network can recognize patterns regardless of their position in the input data, making it more robust to translations and slight spatial variations. Max-pooling reduces CNN overfitting. Discarding non-maximum values reduces feature map spatial dimensions and network parameters and computations.

- Soft Max:

The SoftMax function converts a vector of numbers into a probability distribution with probabilities proportional to their magnitudes. Practical machine learning neural network models use it as an activation function. The SoftMax function is usually applied to the last Dense layer output in a neural network. SoftMax ensures non-negative values by applying the exponential function to each input. Dividing each element by the sum of all exponential values normalizes these values.[18] The SoftMax function outputs a vector of probabilities with

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

each element representing the probability of a class.

- Output:

There are 2 different categories of classes that has been categorized, which includes normal and Abnormal.

B. Proposed Convolutional Neural Network (CNN)

Convolutional neural networks (CNNs) have become prominent in various computer vision tasks and are increasingly being utilized in diverse fields such as radiology. CNNs consist of different components, incorporating fully connected, pooling, and convolutional layers as shown in figure 3. Their goal is to use the backpropagation technique to independently and dynamically learn spatial hierarchies of features.

Figure 3. Proposed CNN model

Image Augmentation:

To enhance the robustness and generalization capabilities of our CNN model, we employed various image augmentation techniques during the training phase. These techniques were implemented using the Keras Image Data Generator class, which provides a convenient way to apply transformations to the training dataset. First, we applied rescaling to normalize the pixel values of the images. This rescaling step helps the model converge faster during training by ensuring consistent input ranges. Next, we incorporated shear transformations by specifying a shear range. Shearing involves shifting parts of the image along a certain axis, resulting in a distorted effect. This augmentation technique helps the model learn to be invariant to changes in orientation. Additionally, we applied zooming transformations to the images by defining a zoom range. This technique randomly magnifies or reduces the size of the image, simulating different perspectives or scales. By incorporating zooming, our model becomes more robust to variations in object size and allows for better generalization. Furthermore, we employed horizontal flipping as an augmentation technique. By randomly flipping images horizontally, we introduced additional variations without changing the object's identity or meaning. This augmentation is particularly useful when the orientation of the object is not crucial for the task at hand, such as in object recognition. By leveraging these image augmentation techniques, we

aimed to increase the diversity and size of our training dataset, thereby improving the model's ability to generalize to unseen data and handle variations in the input images. The augmented dataset was then used to train our CNN model, enabling it to learn robust and discriminative features necessary for accurate classification. Overall, the inclusion of image augmentation techniques in our CNN model's methodology strengthens its ability to handle various real-world scenarios, contributing to the robustness and performance of the research outcomes.

Convolutional layer:

To capture and extract relevant features from the input images, we designed a Convolutional Neural Network (CNN) architecture. The CNN model was implemented using the Keras framework. The first step in constructing the CNN involved initializing the classifier using the Sequential class, which allows us to build a sequential stack of layers. Next, we added the initial convolutional layer. This layer plays a crucial role in detecting local patterns and features in the input images. We used the Convolution2D layer with 32 filters, a kernel size of 1x1, and a ReLU activation function. The input shape was set to (64, 64, 3), indicating that our model expects images of size 64x64 with 3 color channels (RGB).

Max pooling layer:

To reduce the spatial dimensions of the feature maps and extract the most important information, we incorporated MaxPooling layers into our Convolutional Neural Network (CNN) architecture. MaxPooling is a commonly used technique in CNNs for down-sampling feature maps. After each convolutional layer, we applied a MaxPooling layer. This layer operates on local neighborhoods within the feature maps and selects the maximum value from each neighborhood, effectively reducing the dimensionality of the feature maps. In our model, we used a pool size of 2x2 for all the MaxPooling layers, which means the feature maps were divided into non-overlapping regions of size 2x2, and the maximum value within each region was retained. This process resulted in feature maps with reduced spatial dimensions. By incorporating MaxPooling layers, we achieved several benefits. Firstly, it helped to abstract the representation of features, enabling the model to focus on the most salient and discriminative elements within each region. This abstraction facilitates the model's ability to generalize and recognize patterns even with slight variations in the input images. Secondly, MaxPooling reduces the number of parameters in the network, leading to a more efficient and computationally tractable model. The down-sampling achieved by MaxPooling helps to alleviate the risk of overfitting by preventing the model from memorizing specific details and encouraging it to capture more robust and invariant features. Overall, the MaxPooling layers in our CNN model played a crucial role in reducing spatial dimensions, abstracting features, and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

enhancing the model's efficiency and generalization capabilities. By incorporating MaxPooling, our CNN architecture successfully captured hierarchical representations, enabling accurate classification and yielding promising results in our research study. [19]

ReLU:

We utilized the ReLU activation function in the convolutional layers. ReLU is a simple yet effective activation function that introduces non-linearity into the network, allowing it to learn complex and non-linear relationships between the input data and the corresponding features. By applying the ReLU activation function after each convolution operation, we introduced non-linearity into the model, enabling it to learn and capture more complex representations. ReLU has been shown to mitigate the vanishing gradient problem, allowing the network to learn more effectively and preventing the gradients from diminishing as they propagate backward.

Flattening:

A single, long continuous linear vector is created by flattening all of the obtained 2-Dimensional arrays from the pooled feature maps. The completely connected layer receives the flattened matrix as input to categorize the picture.

Softmax:

We incorporated the softmax activation function in the last Dense layer. Softmax is a probabilistic activation function that converts the outputs of the previous layer into a probability distribution over multiple classes. The softmax function takes the inputs and applies the exponential function to each element. It then normalizes the resulting values by dividing each element by the sum of all the exponential values. The output of the softmax function is a vector of probabilities, with each element representing the probability of the corresponding class. Overall, the utilization of the softmax activation function in the final layer enhances the interpretability and performance of our CNN model in multi-class classification tasks.[20] It provides a probabilistic framework for class predictions, enabling us to make informed decisions and assess the certainty of our model's outputs in the context of our research study.

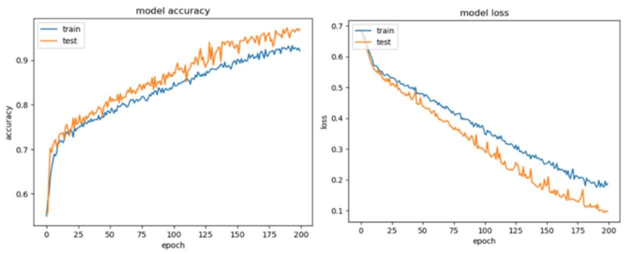

Figure 4. (i)Graph of number of epochs vs accuracy of CNN model, (ii)Graph of number of epochs vs loss of CNN model

(i) A graph in figure 4(i), the number of epochs and CNN model accuracy:

A Convolutional Neural Network (CNN) model's performance is shown in the graph across a number of epochs. The training data performed a specified number of full runs through the model at each epoch, which is represented by the number of epochs on the x-axis. The accuracy of the CNN model is shown by the y-axis, which shows the proportion of cases that the model correctly categorized.

(ii) CNN Model Number of Epochs vs. Loss graph shown in figure 4(ii)

The graph shows how a Convolutional Neural Network (CNN) model degrades with more epochs. The training data's total number of complete passes during the model-training process is shown on the x-axis as the number of epochs. The CNN model's loss, which is measured as the difference between the model's actual results and its anticipated results, is shown on the y-axis.

C. Pre-Trained models

VGG16

VGG is an object identification and classification method that achieves an accuracy rate of 92.7% when categorizing 1000 pictures into 1000 different categories. Its 16 weighted layers—out of a total of 21 layers—are denoted by the term "VGG16". There are 13 convolutional layers, 5 Max Pooling layers, and 3 Dense layers included in this ensemble. VGG16 requires a 224x224 input tensor with three RGB channels. There are 64 filters in the Conv-1 Layer, 128 filters in the Conv-2 Layer, 256 filters in the Conv-3 Layer, and 512 filters in the Conv-4 and Conv-5 Layers. Everywhere in the design of VGG16, convolution and max pool layers are distributed equally.

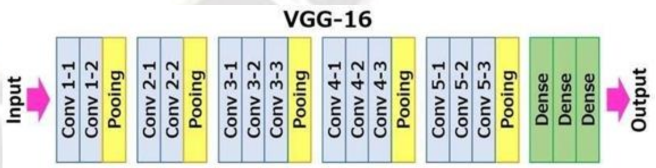

Figure 5. Architecture of VGG-16 model

As shown in figure 5, The network receives a 224x224x3 image. The first two layers have 64 channels and 3x3 filters with the same padding. A max pooling layer with stride (2, 2) follows.

Convolutional layers with 128 filters and (3,3) filter size follow. Another max pooling layer with stride (2, 2) follows. Next, two convolution layers with 256 filters and 3x3 filter sizes are distributed.[21] A max pooling layer follows two sets of three convolution layers. Each set has 512 filters with padding of size 3x3. These convolution and max pooling layers use 3x3 filters instead of 7x7 filters like AlexNet. Some layers adjust input

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

channels using 1x1 pixels. 1-pixel padding after each convolution layer preserves image spatiality. We get a 7x7x512 feature map by adding a convolution and max pooling layer to the stack. Flattening yields a 1x25088 feature vector. The three linked layers generate a 1x4096 vector from the flattened feature vector. The second layer then creates a 1x4096 vector. The Softmax function divides input into 1000 categories using a 1x1000 vector from the third layer. Each hidden layer's activation function is ReLU, accelerating learning and reducing errors.

ResNet-50

ResNet50 is a convolutional neural network that consists of 50 layers. ResNet, also known as Residual Networks, forms the basis for various computer vision applications. It is capable of training extremely deep neural networks with 150 or more layers.

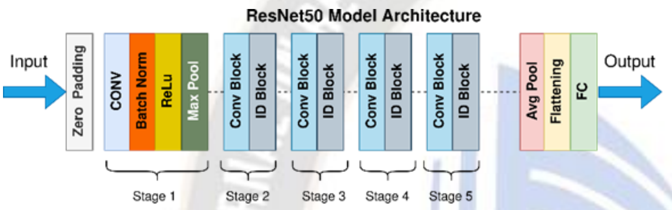

Figure 6. Architecture of ResNet50

The ResNet architecture shown in figure 6 adheres to two key design principles. Firstly, each layer contains an equal number of filters regardless of the output feature map size. Secondly, to maintain consistent time complexitythe number of filters doubles if the size of the feature map is cut in half. These guidelines are followed by the ResNet-50 design, with one major deviation. It includes the "bottleneck" building block, which uses 1x1 convolutions to cut down on the amount of parameters and matrix multiplications. Each layer's training becomes quicker as a result. Instead of using a two-level stack, the 50-layer ResNet utilizes three layers. The components of the ResNet-50 design are as follows (listed below).

- Convolutions with 7x7 kernels and 64 additional kernels, with stride size of 2.Maximum layer of pooling with stride size 2.
- Nine additional layers, including a convolutional layer with 3x3 kernels and 64 cores, a layer with 1x1 kernels and 64 cores, and a third layer with 1x1 kernels and 256 cores. This pattern is repeated three times.
- Four sets of layers, each consisting of a convolutional layer with 1x1 kernels and 512 cores, a convolutional layer with 3x3 kernels and 128 cores, and another convolutional layer with 1x1 kernels and 128 cores.
- Over 18 layers, including a layer with 1x1 kernels and 256 cores, a layer with 2x2 kernels and 256 cores, and a layer with 1x1 kernels and 1024 cores. This pattern is repeated six times.
- The number of filters doubles if the size of the feature map is cut in half. These guidelines are followed by the ResNet-50 design, with one major deviation. It includes the "bottleneck" building block, which uses 1x1 convolutions to cut down on the amount of parameters and matrix multiplications. Each layer's training becomes quicker as a result.

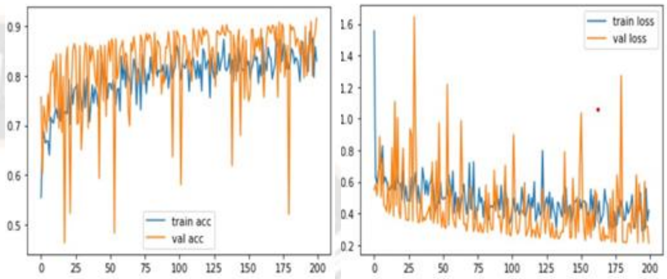

Figure 7. (i)Graph of number of epochs vs accuracy of the ResNet-50, (ii)Graph of number of epochs vs loss of the ResNet-50.

(i) Graph of the ResNet-50 model's shown in figure 7 (i) accuracy vs number of epochs:

The graph displays the ResNet-50 model's performance throughout several epochs. The model goes through a certain number of full iterations during the training phase, which is represented by the number of epochs on the x-axis. The proportion of cases that the ResNet-50 model correctly predicted is shown on the y-axis as the model's accuracy.

(ii) ResNet-50 Model Number of Epochs vs. Loss Graph shown in figure 7 (ii)

As the number of epochs rises, the graph shows the ResNet-50 model's loss. The model's training phase involved a certain number of full iterations, which are represented by the number of epochs on the x-axis. The loss of the ResNet-50 model, which is a measurement of the discrepancy between the model's expected and actual outputs, is shown on the y-axis.

Inception V3

Inception-v3 is a significant architecture shown in figure 8, from the Inception family of convolutional networks, which are often employed in computer vision problems. It includes a number of upgrades, including Label Smoothing, factorized 7x7 convolutions, and an extra classifier to spread label information farther across the network. Convolutions, average pooling, dropouts, concatenations, max pooling, and fully linked layers are some of the symmetric and asymmetric building elements included in the model. The activation inputs are batch-normalized, and SoftMax is employed to compute the loss.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
***Article Received**: 25 April 2023 **Revised:** 12 June 2023 **Accepted:** 28 June 2023*

Inception V3, which was released in 2015, contains 42 layers and is less error-prone than its forerunners. The model has been improved with the use of auxiliary classifiers, factorization into smaller convolutions, spatial factorization into asymmetric convolutions, and effective grid size reduction.[22]

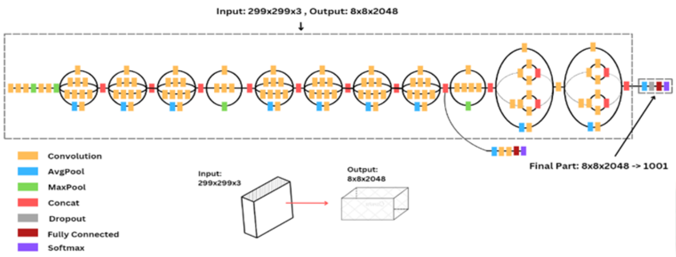

Figure 8. Architecture of Inception V3

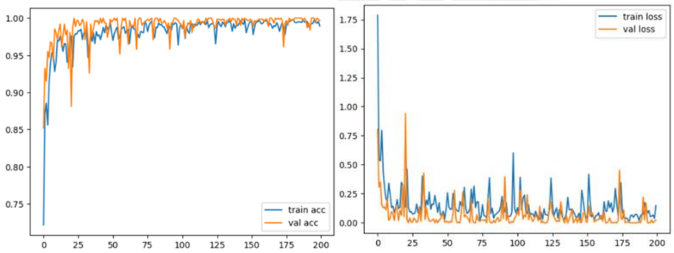

Figure 9. (i)Graph of number of epochs vs accuracy of the Inception-V3, (ii)Graph of number of epochs vs loss of the Inception-V3.

(i) Graph showing Number of Epochs vs. Inception-V3 Model shown in figure 9 (i) Accuracy:

The Inception-V3 model's performance throughout several epochs is shown in the graph. The model goes through a certain number of full iterations during the training phase, which is represented by the number of epochs on the x-axis. The percentage of cases that the Inception-V3 model correctly predicted is shown on the y-axis as the model's accuracy.

(ii) The Inception-V3 Model's Number of Epochs vs. Loss Graph shown in figure 9 (iI)

The graph shows how the Inception-V3 model degrades as the number of epochs rises. The model's training phase involved a certain number of full iterations, which are represented by the number of epochs on the x-axis. The loss of the Inception-V3 model, which is measured as the difference between the model's anticipated and actual outputs, is shown on the y-axis.

IV. RESULTS AND DISCUSSION

The dataset consists of 2,733 Chest X-ray images that 8 bits each. There are 2 defined classes of Chest X-rays presented in this paper with 1 class of abnormal chest X-rays and the other class is Normal Chest X-ray (i.e., the null case). For detection test image is taken then pre-processing is done which includes resizing, filtering and threshold. The training set consists of all the Abnormal Chest X-rays. Then depending on maximum matched trained images, test images classified in either of two classes. The class is decided according to index number of trained images. In this way the model is trained, and the image can be classified into Normal and Abnormal Chest X-ray.

The pre-trained models were designed and trained on the VGG16, ResNet50 and Inception V3 but directly using them for transfer technique for the Chest X-rays Dicom dataset is not possible. This is due to the fact that the pre-trained models were developed on RGB images, which have quite different physical properties than DICOM images. So, we converted the dataset from DICOM format to JPEG format, and trained it on our CNN model as well as other pre-trained model. The validation accuracy obtained with the CNN model is 96.95%, InceptionV3 is 95.34%, VGG16 is 94.81% and Resnet-50 is 90.43%. These models will be utilized to perform automatic classification of Abnormal and Normal Chest X-rays. The comparison chart of the CNN model and other pre-trained models is given in Table 1. Figure 10. Shows the graph of no. of epochs vs accuracy of the CNN model

After 200 iterations with the SGD Optimizer, the CNN model has training accuracy of 96.22% and validation accuracy of 96.95%. The performance may be shown by graphing the accuracy graph and the number of epochs. The photographs that go with it show the outcomes. The model had an F1 score of 0.533, accuracy of 0.5334, and recall of 0.538.

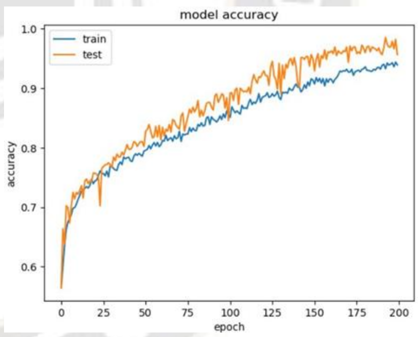

Figure 10. Graph of No. of epochs vs accuracy of the CNN model

The table 1 suggests that the CNN Model and Inception V3 are the most efficient transfer learning techniques for classification of this Chest X-ray dataset. Further analysis of precision, recall, and F1 Score values indicate that CNN model is the optimal choice.

Table 1. Comparision Table

Transfer Learning Techniques Parameters	CNN Model	InceptionV3	VGG16	ResNet-50
No. of Epochs	200	200	200	200
Batch Size	32	32	32	32
Optimizer	SGD	Adam	Adam	Adam
Validation	96.95%	95.34%	94.81%	

Accuracy				91.64%
Training Accuracy	96.22%	98.48%	93.26%	83.04%
F1 score	0.69	0.62	0.59	0.52
Recall	0.638	0.59	0.57	0.55
Precision score	0.733	0.67	0.61	0.56

From table 1, we can compare the accuracy and parameters of Inception V3 and VGG16 pre-trained utilized in different research papers with our pre-trained models, who have used different number of unlabeled and labeled dataset used for training their models. In order to improve the accuracy of the CNN model, we have converted a real-time DICOM Dataset into JPEG format. The labeled data comprises two classes: Normal and Abnormal. The CNN model is designed to self-learn the patterns and repetitions of abnormality present in the dataset. This training process equips the model with the ability to distinguish between normal and abnormal features when presented with an unlabeled chest X-ray image.

V. CONCLUSION

The proposed approach of using CNNs for the classification of chest X-ray images in a medical decision support system has proven to be effective and accurate. The customized CNN model designed for this project outperformed the pre-trained models with a validation accuracy of 96.95%, as shown in table 1, making it a valuable tool for medical professionals in the diagnosis and treatment of chest-related diseases in patients. The results obtained from this study show that deep learning techniques such as CNNs can significantly aid radiologists in medical decision-making processes, especially in situations where the availability of experts is limited. Furthermore, the use of original labeled dataset with 1151 normal and 1376 abnormal chest X-rays in DICOM format ensured that the CNN model was well-equipped to handle real-world scenarios and provide accurate diagnoses. As the shortage of radiologists continues to be a growing concern in the medical field, the implementation of machine learning and artificial intelligence techniques can potentially help bridge the gap and improve patient outcomes. Overall, this project highlights the potential of using CNNs for medical decision support systems and emphasizes the importance of using original data for training and testing models in the field of medical imaging.

REFERENCES

[1] Rahmat, Taufik & Ismail, Azlan & Aliman, Sharifah. (2018). CHEST X-RAYS IMAGE CLASSIFICATION IN MEDICAL IMAGE ANALYSIS. Applied Medical Informatics. 40. 63-73.

[2] Apostolopoulos, I.D., Mpesiana, T.A. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43, 635–640 (2020). https://doi.org/10.1007/s13246-020-00865-4

[3] Puttagunta, M., Ravi, S. Medical image analysis based on deep learning approach. Multimed Tools Appl 80, 24365–24398 (2021). https://doi.org/10.1007/s11042-021-10707-4

[4] Kim S, Rim B, Choi S, Lee A, Min S, Hong M. Deep Learning in Multi-Class Lung Diseases' Classification on Chest X-ray Images. Diagnostics (Basel). 2022 Apr 6;12(4):915. doi: 10.3390/diagnostics12040915. PMID: 35453963; PMCID: PMC9025806.

[5] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen, "Medical image classification with convolutional neural network," 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, 2014, pp. 844-848, doi: 10.1109/ICARCV.2014.7064414.

[6] Yang, Ruixin & Yu, Yingyan. (2021). Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis. Frontiers in Oncology. 11. 638182. 10.3389/fonc.2021.638182.

[7] Akshay Kumar S, Fida Mohammed Khaleel, Karthika A J, Gayathri J L, Allen Saju N, Sujarani M S, 2021, Medical Image Classification and Cancer Detection using Deep Convolutional Neural Networks, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCREIS – 2021 (Volume 09 – Issue 13),

[8] Sara Dilshad, Nikhil Singh, M. Atif, Atif Hanif, Nafeesah Yaqub, W.A. Farooq, Hijaz Ahmad, Yu-ming Chu, Muhammad Tamoor Masood, Automated image classification of chest X-rays of COVID-19 using deep transfer learning, Results in Physics, Volume 28, 2021, 104529, ISSN 2211-3797, https://doi.org/10.1016/j.rinp.2021.104529.

[9] Yadav, S.S., Jadhav, S.M. Deep convolutional neural network based medical image classification for disease diagnosis. J Big Data 6, 113 (2019). https://doi.org/10.1186/s40537-019-0276-2

[10] Pham, Hieu, Dung Do and Ha Q. Nguyen. “DICOM Imaging Router: An Open Deep Learning Framework for Classification of Body Parts from DICOM X-ray Scans.” ArXiv abs/2108.06490 (2021): n. pag.

[11] Yan, Zhennan & Zhan, Yiqiang & Peng, Zhigang & Liao, Shu & Shinagawa, Yoshihisa & Metaxas, Dimitris & Zhou, Xiang. (2015). Bodypart Recognition Using Multi-stage Deep Learning. Information processing in medical imaging : proceedings of the ... conference. 24. 449-61. 10.1007/978-3-319-19992-4_35.

[12] H. Sun, Z. Liu, G. Wang, W. Lian and J. Ma, "Intelligent Analysis of Medical Big Data Based on Deep Learning," in IEEE Access, vol. 7, pp. 142022-142037, 2019, doi: 10.1109/ACCESS.2019.2942937.

[13] A. Bah and M. Davud, "Analysis of Breast Cancer Classification with Machine Learning based Algorithms," 2022 2nd International Conference on Computing and Machine Intelligence (ICMI), Istanbul, Turkey, 2022, pp. 1-4, doi: 10.1109/ICMI55296.2022.9873696.

[14] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan and A. Mittal, "Pneumonia Detection Using CNN based Feature Extraction," 2019 IEEE International Conference on Electrical, Computer

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 8s
DOI: https://doi.org/10.17762/ijritcc.v11i8s.7193
Article Received: 25 April 2023 Revised: 12 June 2023 Accepted: 28 June 2023

and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-7, doi: 10.1109/ICECCT.2019.8869364.

[15] Masadeh, Mahmoud & Masadeh, Ayah & Shorman, Omar & Khasawneh, Falak & Masadeh, Mahmoud. (2022). An Efficient Machine Learning-Based COVID-19 Identification Utilizing Chest X-Ray Images. IAES International Journal of Artificial Intelligence (IJ-AI). 11. 1-11. 10.11591/ijai.v11.i1.pp356-366.

[16] Juan Garcia, Guðmundsdóttir Anna, Johansson Anna, Maria Jansen, Anna Wagner. Optimal Decision Making in Supply Chain Management using Machine Learning. Kuwait Journal of Machine Learning, 2(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/209

[17] Hwa, Stefanus Kieu Tao, Mohd. Hanafi Ahmad Hijazi, Abdullah Bade, Razali Yaakob and Mohammad Saffree Jeffree. “Ensemble deep learning for tuberculosis detection using chest X-ray and canny edge detected images.” IAES International Journal of Artificial Intelligence (IJ-AI) (2019): n. pag.

[18] Mandeel, Thulfiqar & Awad, Salah & Naji, Shama. (2022). Pneumonia binary classification using multi-scale feature classification network on chest x-ray images. IAES International Journal of Artificial Intelligence. 11. 1469-1477. 10.11591/ijai.v11.i4.pp1469-1477.

[19] Hwa, Stefanus & Bade, Abdullah & Hijazi, Mohd & Jeffree, Mohammad. (2020). Tuberculosis detection using deep learning and contrastenhanced canny edge detected X-Ray images. IAES International Journal of Artificial Intelligence (IJ-AI). 9. 713. 10.11591/ijai.v9.i4.pp713-720.

[20] F.-F. Li and A. Karpathy, "Convolutional Neural Networks for Visual Recognition," Stanford University, [Online]. Available: http://cs231n.github.io/. [Accessed: May 7, 2023].

[21] F. Chollet, "Deep Learning with Python," Manning Publications, 2017.

[22] A. Rosebrock, "Deep Learning for Computer Vision with Python," PyImageSearch, 2018.

[23] A. Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow," O'Reilly Media, 2019.