DETECTION OF COVID 19, PNEUMONIA, TUBERCULOSIS USING CONVOLUTIONAL NEURAL NETWORKS AND DEEP LEARNING ARCHITECTURES

A PROJECT REPORT

Submitted in partial fulfillment of the requirements for the award of the internship

By

Anil Sinthu (22VV5A1273)

Chandra Sai Shankar Karra (21VV1A1231)

(Students of JNTU-GV COLLEGE OF ENGINEERING VIZIANAGARAM (A))

Supervisor:

Dr. Mohammad Farukh Hashmi

Assistant Professor

Department of Electronics and Communication Engineering

Department of Electronics and Communication Engineering

National Institute of Technology, Warangal

2023-24

ABSTRACT:

This paper offers a comprehensive assessment of the possibilities, applications, and technologies of employing convolutional neural networks and deep learning to identify pneumonia. To choose pertinent research sources, this study employs a synthetic methodology. The report compares and evaluates several computer-aided strategies for lung illness identification and provides an improved model for pneumonia detection that we will apply in our next research. The most widely used databases, including Google Scholar, were searched for the study. A total of thirty publications were gathered, of which fifteen were thoroughly examined about the subject matter under investigation. Additionally, the review study highlights the importance of pneumonia, TB, normal, and COVID-19. Pneumonia is a bacterial illness that can kill humans; it usually affects one or both

Keywords: Chest X-ray pictures, pneumonia, covid19, tuberculosis, deep learning, transfer learning, convolution neural network (CNN), and computer-aided diagnosis.

1. INTRODUCTION:

Indoor air pollution is one of the main causes of pneumonia in children. Pneumonia is a severe acute respiratory illness that affects many people in age groups at the extremes of the average human lifetime. The cause is high infectious organism attack rates [1]. In the US alone, pneumonia causes over a million hospital admissions as well as about 50,000 adult fatalities annually [2]. Medical personnel use a range of techniques, including physical examinations, medical histories, clinical investigations (sputum or blood tests), chest X-rays, and other imaging modalities, to diagnose pneumonia in hospital patients. William Osler used the term "the captain of the men of death" to describe pneumonia in the 1800s.

Machine learning research has shifted its attention to computer-aided designs (CAD) in recent years. Benefits of CAD include enhanced diagnostic accuracy, decreased workload pressure, and increased inter- and intra-reader variability [3]. Convolutional neural networks (CNNs) are employed to solve challenging image-driven pattern recognition problems because they can identify both linear and nonlinear patterns. CNNs draw inspiration from the visual cortex of the brain [4]. Convolutional Neural Networks (CNNs) have demonstrated their capacity to extract relevant properties from pictures when used as Deep Learning (DL) models [5,6].

VGG16, VGG19, ResNet50 and its variations, Inception-v3 variations, MobileNetV3, DenseNet121, and EfficientNetb0 are transfer learning models available on Keras. Pneumonia patients have symptoms in the chest cavity such as fluids filling the lungs' air sacs, and their radiological picture appears brighter. The goal of the work is to create CNN models from keras that have high validation accuracy, recall, and F1 score, and can identify pneumonic patients from chest X-rays. The effectiveness of deep learning approaches has been demonstrated, and their accuracy in predicting diseases is comparable to that of an average radiology [7]. Numerous tests are available for diagnosing pneumonia, including computed tomography of the lungs, needle biopsy of the lung, chest MRI, chest X-ray, and chest ultrasound.

The following are the contributions.

1) A CNN model is trained on photos to recognize a variety of illnesses, such as pneumonia, Covid 19, normal, and tuberculosis.
2) Differences in the model's mutual accuracy result from modifications. The CNN model's layers are changed to improve the model's accuracy.
3) Model assessment measures such as f1 score, recall, and precision, as well as model accuracy and loss, improve with the optimizer (Adam) selection, pre-processing and data augmentation modifications, and parameter fine-tuning (rotation, flip, learning rate, epochs, and batch size).
4) To assess the model layers and the size and shape of the picture targets used in model training and testing, use the confusion matrix and classification matrix.

2. LITERATURE SURVEY:

In the field of medical diagnostics, the problem of categorizing chest X-ray images into distinct classifications has been thoroughly investigated. Scholars have attempted to employ many techniques to reduce dimensionality in [8]. The scientists evaluated the artificial neural network [9] to identify lung conditions such as lung cancer, pneumonia, and TB. The study of medical image categorization has drawn more attention lately as machine learning (ML) techniques for the identification of thoracic illnesses are being investigated. Lakhani and Sundaram (2017) [10] developed a method based on the design of two different CNNs, Alex Net and Google Net, for diagnosing pulmonary TB. This dataset was recently explored by Pranav Rajpurkar, Jeremy Irvin, et al. (2017) [11] to see if it could predict pneumonia more reliably than radiologists. Their model, which they named Chex Net, uses a DenseNet-121-layer architecture to identify every one of the 14 illnesses from the dataset's enormous 112,200 photos.

Researchers describe a method for detecting lung cancer [12] using chest X-ray imaging. In the study [13], a modified CNN with meticulous regulation was used to identify X-ray thorax sickness in the chest. Anthimopoulos et al. [14] developed a CNN model to identify interstitial lung disease patterns. The dataset utilized for training comprises seven classes, which are represented by 14,696 photos. Their model attained 85.5% accuracy. Krizhevsky et al. [15] showed that deep CNN models, with a top 5 error percentage of 17%, may yield creative results on difficult datasets. The dataset that was chosen was ImageNet. The total population inside their network is sixty million. parameters and includes five convolutional layers along with max-pooling layers. Three completely linked layers were used to get the best results. Benjamin Antin et al. (2017) [16] analyzed the dataset [11] after the Chex Net [24] model was developed, and they proposed a logistic regression model for pneumonia detection.

Deep learning-based methods have been used in several disciplines in the past [17–21]. A multitude of biological image identification techniques have previously been put out by different writers. The challenges and future potential of medical image processing were discussed by Razaak [22]. To accurately identify and classify the distinct types of chest X-rays, CNNs work based on the characteristics that are collected by the various layers. Roth et al. [23] created an algorithm for lymph node identification in computed tomography data using deep CNN and a 2.5D resampling method. The chest X-ray pictures are subjected to an assessment process known as scan line optimization, which removes any other body parts, to avoid any diagnostic mistakes. Rajaraman et al. [24] published a visualization technique for locating the region of interest in chest X-rays. Rahib et al. [25] and Okeke et al. [26] use the CNN model to categorize pneumonia patients.

In this work, we started with two class datasets and combined them with a TB dataset, and a COVID-19 dataset and made this dataset with four classes called Pneumonia, Normal, Covid-19, and Tuberculosis. The CNN model was utilized to improve accuracy.

3. METHODS:

3.1. CNN:

A type of deep neural network called a convolutional neural network (CNN) is used to analyze visual images. It has many hidden layers, input layers, and output layers [27]. A CNN-based machine learning system is used to classify the image. One type of deep learning neural network is CNN [28]. We use a total of ten layers in our CNN model: two thickness layers, one smoothing layer, one output layer, three MaxPooling2D layers, and three Convolution2D layers.

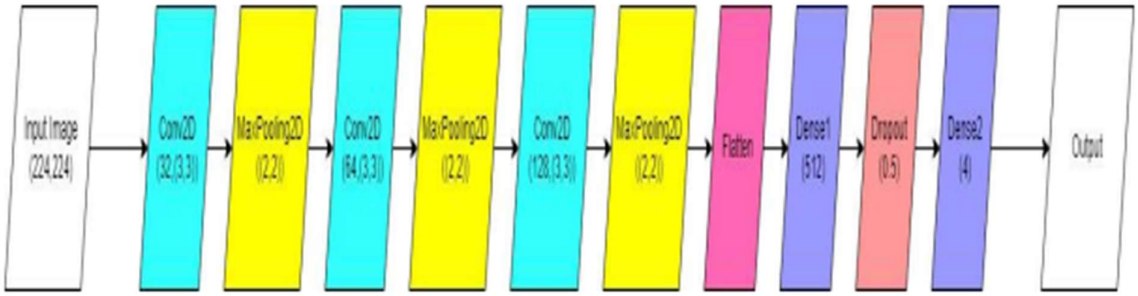

Figure 1: **CNN** Sample Architecture

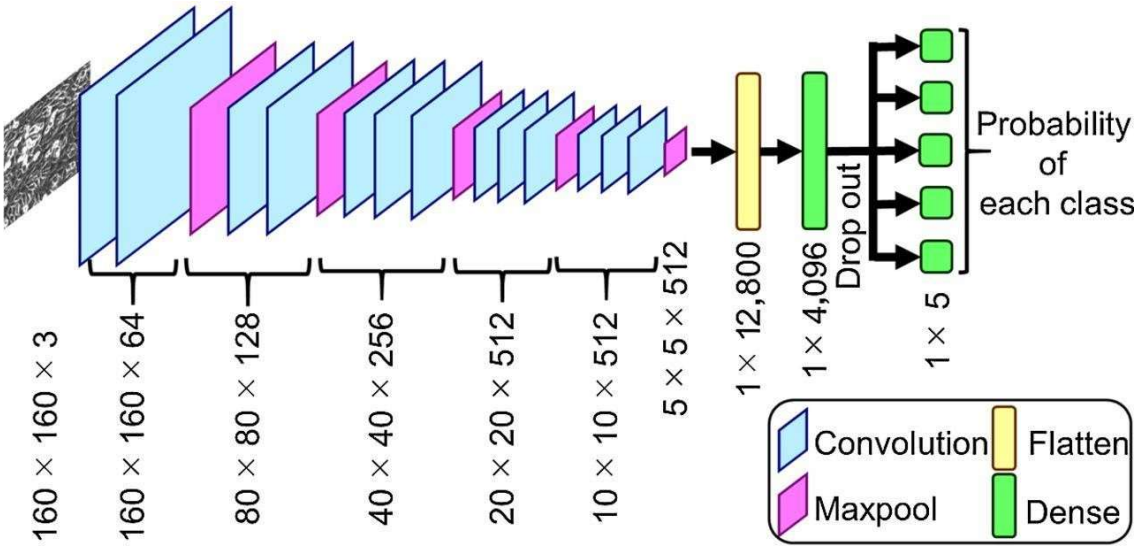

Figure 2: **CNN** Model Architecture

3.1.1. Convolution Layer:

A picture input is transformed into a matrix form. A feature map is produced when the input matrix and a feature detector, filter, or kernel of dimension 3X3 are subjected to the convolution operation [29]. The image's size is decreased during this operation, simplifying image processing. This also results in information loss, but the feature detector keeps the essential components of the picture [30].

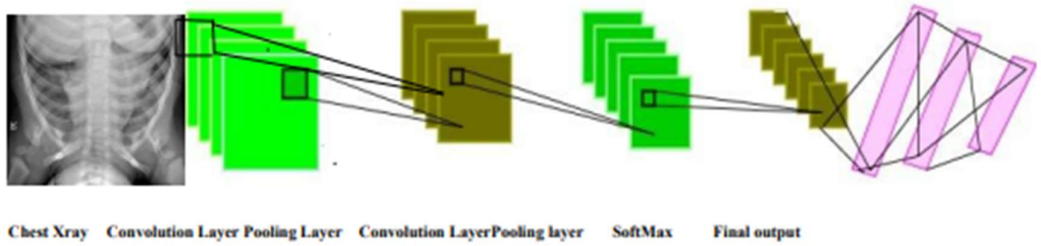

Figure 3: Convolutional neural network consisting of convolution and pooling layers and fully connected SoftMax layers at the end to give the final prediction

3.1.2. Max-Pooling Layer:

The pooling layer's objective is to further down-sample the input image. To put it another way, to make the input image smaller [31]. As a result of the image's fewer parameters, the computational complexity is decreased. The models employ max-pooling and average-pooling subsampling techniques. A sample-based discretization technique is called max-pooling. Each feature map is processed by the dimension 2X2 pooling layer, which uses the 'MAX' function to scale the dimensionality. The highest pixel value from the image window that the feature detector is currently covering is chosen using max pooling [32]. The pooling layer is a feature extraction operation, which is typically applied in a Convolutional Neural Network (CNN).

3.2. TRANSFER LEARNING:

Transfer learning is a powerful deep-learning technique. Applying a model that has already been trained to a new problem is known as transfer learning. Transfer learning is a technique whereby pre-trained machine learning model data is migrated to a different but related model. Transfer learning is a concept that may be used for a smaller dataset by using a trained model from a larger dataset, such as ImageNet, as shown in Figure 2 [33]. In machine learning (ML), transfer learning is used. In Transfer Learning, a machine makes better generalizations about one job by using the knowledge it has learned from an earlier one.

In this case, X is a specific learning sample ($X = x1, \cdots, xn, \in \chi$), xi is the ith term vector corresponding to certain texts, and χ is the space of all term vectors.
The task T for a given domain D is defined with the following parameters:
$T = \{\gamma, P(Y \mid X)\} = \{\gamma, Y = \{y1, \cdots, yn\}, yi \in \gamma$ **(1)**

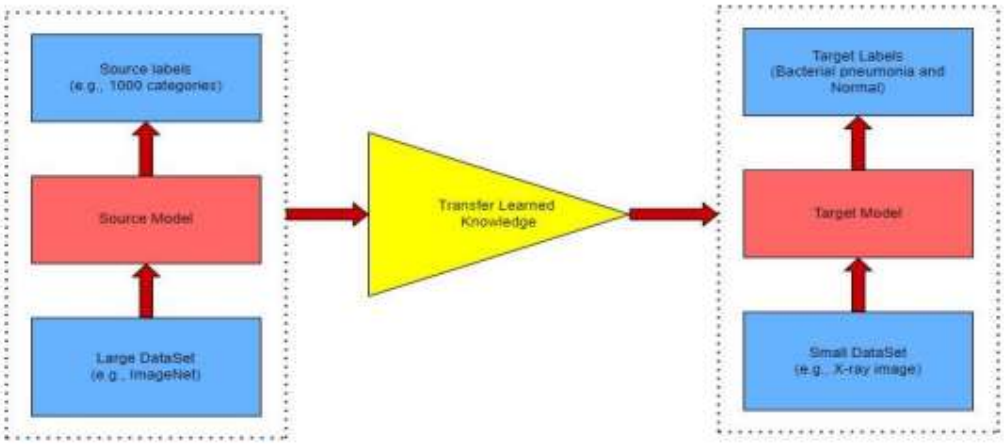

Figure 4: Concept of Transfer Learning

In Summary, The Neural Network consists of three convolution layers and each layer consists of a pooling layer and a dropout layer. Max Pooling of pool size of (2, 2), dropout of 0.5, and "ReLU" as activation function is used. The model is then flattened, and a dense layer is added with 512 pixels along with the activation function "ReLU". Afterward, a dropout layer of 0.5 is added again to prevent overfitting by excluding the least wanted features. This is followed by a dense layer with an activation function "softmax" to fetch the probabilistic distribution on the output. Adam optimizer (Adaptive Moment Estimation) is used and a loss of "categorical cross entropy" for binary classification and metrics of "accuracy." Adam optimizer is used as it is considered to provide a better result than any other optimization algorithm. Also, it requires very few parameters and has a faster computation time than others with a smaller learning rate, say 0.001. This faster pace is achieved by the following equation:

$$Wt + 1 = Wt - \alpha mt$$

Where, $mt = \beta mt\text{-}1 + (1\text{-}\beta)\,[\delta L / \delta Wt]$ **(2)**

4. MATERIALS:

4.1. EXPERIMENTAL DATASET:

The combination of the Covid-19 chest x-ray, tuberculosis, and pneumonia datasets is the dataset that we used. The datasets we downloaded from Kaggle include Chest X-ray Images (Pneumonia) [34], Chest X-ray (Covid-19 & Pneumonia), Curated COVID-19 Chest X-Ray Dataset, and Tuberculosis (TB) Chest X-ray Cleaned Database.

Dataset Type	No of Images
Train	6326
Test	771
Val	38

Testing Data	
Dataset Type	No of Images
Pneumonia	390
Normal	234
Covid 19	106
Tuberculosis	41

Validation Data	
Dataset Type	No of Images
Pneumonia	8
Normal	8
Covid 19	10
Tuberculosis	12

Training Data	
Dataset Type	No of Images
Pneumonia	3875
Normal	1341
Covid 19	460
Tuberculosis	650

Figure 5: Experimental Dataset Details - Train, Test, validation datasets

The base folders are named Test, Train, and Valid, while the subfolders of each folder are named Covid19, Normal, Pneumonia, and Tuberculosis. Chest x-ray images are classified (Categorical data) and placed separately in Covid19, Normal, Pneumonia, and Tuberculosis as subfolders of Test, Train, and Valid.

Category	Training Set	Validation Set	Testing Set
Pneumonia	3875	8	390
Normal	1341	8	234
Covid19	460	10	106
Tuberculosis	650	12	41
Total	6326	38	771
Percentage	89	53	10

Table 1: Description of Experimental Dataset

7135 chest x-ray pictures from the Covid19, Normal, Pneumonia and Tuberculosis classes are included in our collection. The Train Dataset contains 6326 images, the Valid Dataset has 38 images, and the Test dataset contains 771 images.

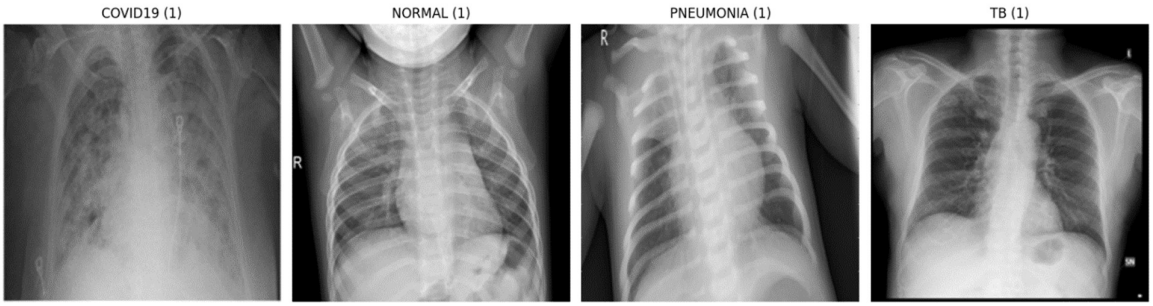

Figure 6: Sample Dataset

5. PROPOSED METHODOLOGY:

This study suggests the best method for determining pneumonia using chest X-rays. Methods for augmenting data were applied to grow the little dataset. The predictions from these models were then combined with a weighted classifier to determine the final forecast. The whole block diagram of the recommended method is shown in Figure 7.

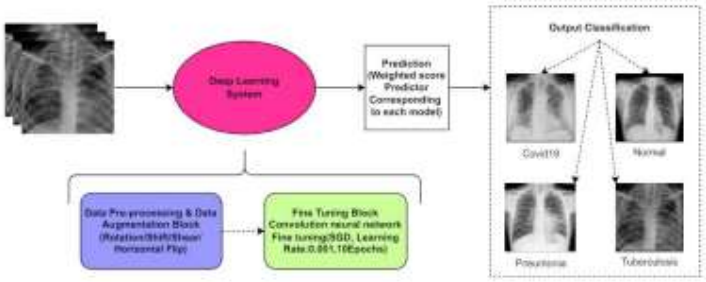

Figure 7: Block diagram of the proposed methodology

5.1. PRE-PROCESSING & DATA AUGMENTATION:

Resizing the X-ray images was a crucial step in the data preprocessing process because different algorithms required varied image input sizes. We set the model's input size to 224,224,3, where 224,224 is the image form and 3 is the number of channels (R, G, and B). Thus, we changed the form of our input image to 224,224.

```
/content/drive/My Drive/CDataset/train/PNEUMONIA/person60_bacteria_286.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person608_virus_1175.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person60_bacteria_287.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person617_bacteria_2488.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person615_virus_1184.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person60_bacteria_285.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person619_bacteria_2491.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person611_bacteria_2476.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person610_bacteria_2475.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person610_virus_1177.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person614_bacteria_2481.jpeg
/content/drive/My Drive/CDataset/train/PNEUMONIA/person612_virus_1179.jpeg
```

Figure 8(a): Preprocessing of Train Dataset

```
/content/drive/My Drive/CDataset/val/COVID19/COVID19(566).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(571).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(567).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(572).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(568).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(569).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(570).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(573).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(574).jpg
/content/drive/My Drive/CDataset/val/COVID19/COVID19(575).jpg
```

Figure 8(b): Preprocessing of Validation Dataset:

```
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-661.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-662.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-660.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-665.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-671.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-664.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-667.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-668.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-666.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-670.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-669.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-663.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-686.png
/content/drive/My Drive/CDataset/test/TURBERCULOSIS/Tuberculosis-682.png
```

Figure 8(c): Preprocessing of Testing Dataset

Data augmentation makes better use of already-existing data to address this issue. To solve this problem, data augmentation maximizes the usage of pre-existing data. It keeps the model from overfitting this dataset and contributes to the current training dataset's expansion.

The photos that follow different enhancing techniques are shown below (Figure 9). Only one of these techniques was used to create the enhanced picture.

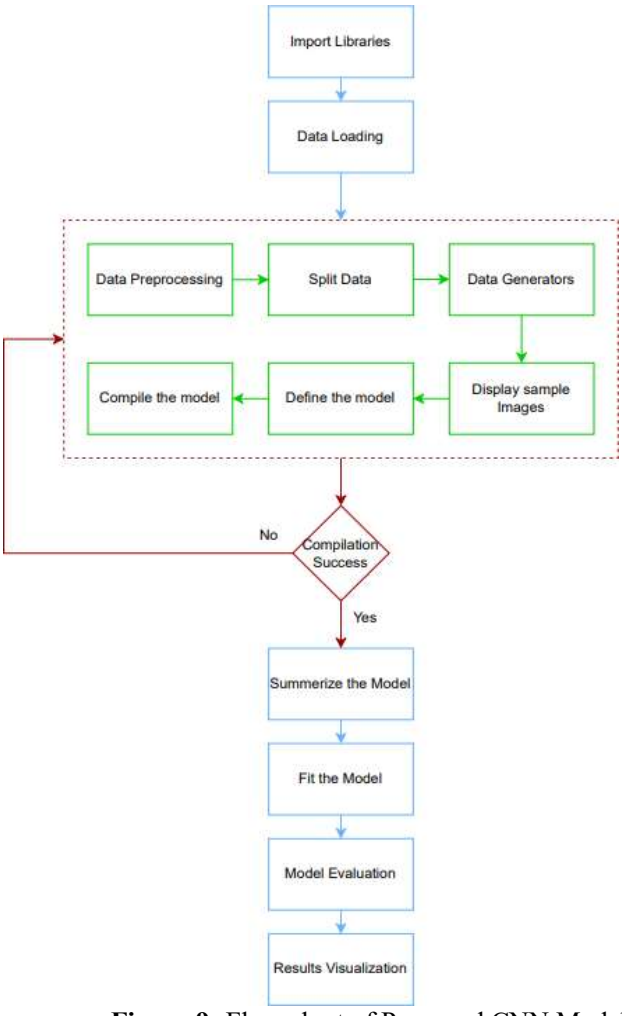

Figure 9: Flow chart of Proposed CNN Model

6. EXPERIMENTAL SETUP:

This section presents the experiments and assessment methods utilized in the research to assess the suggested model's efficacy. Utilized was the chest X-ray image collection that was suggested. The pre-trained architectures were loaded onto the ImageNet Dataset using the open-source Keras deep learning framework, which was then tuned for the given job using TensorFlow as the backend. A standard PC equipped with an Intel Pentium Silver N5000 CPU, 8 GB RAM, and a standard Intel UHD Graphics 605 GPU managed all the processing work.
We used the dataset of 2 classes Pneumonia and Normal and 2 more classes covid19 and Tuberculosis then it was made as a 4-class dataset [44]. In this dataset, we have the base folders named Test, Train, and Valid, while the subfolders of each folder are named Covid19, Normal, Pneumonia, and Tuberculosis. Chest x-ray images are classified (Categorical data) and placed separately in Covid19, Normal, Pneumonia, and Tuberculosis as subfolders of Test, Train, and Valid. 7135 chest x-ray pictures from the Covid19, Normal, Pneumonia and Tuberculosis classes are included in our collection. The Train Dataset contains 6326 images, the Valid Dataset has 38 images, and the Test dataset contains 771 images.

Parameters	**Values**
Number of Classes	4
Image Size	(224,224)
Channels	3
Batch Size	256
Epochs	20
Optimizer	Adam
Learning Rate	0.001
Activation Function	‘ReLu’
Loss Function	categorical_crossentropy

Table 2: Experimental setup of Proposed CNN Model

7. EXPERIMENT RESULTS AND DISCUSSION:

A collection of 16 images from the training dataset is shown, along with the names of their classes. It initially pulls data from the training generator, including class names and related indices. It then obtains an album of pictures together with their captions. It displays each of the first 16 photographs in this batch in a 4x4 grid and normalizes the pixel values for better viewing. Each image's title is set to the name of its class, which is found in the label array by showing the class with the highest probability. For a more polished appearance, the axes are off. The plot is finally shown.

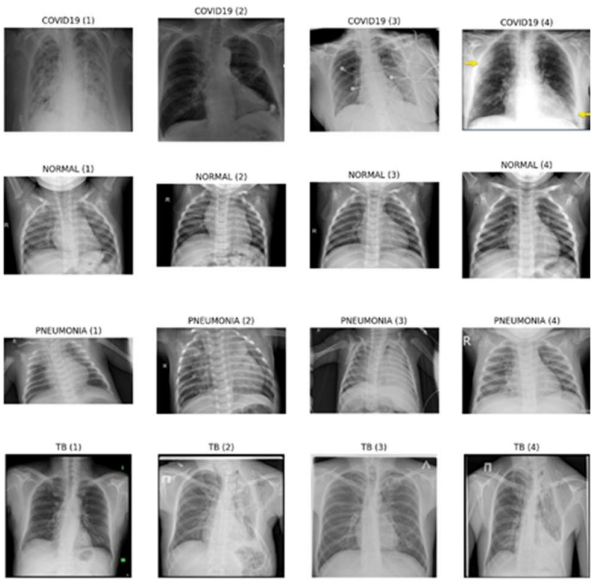

Figure 10: Images of different diseases.

RESULTS:

We also worked on existing models to get more knowledge and to understand models clearly. After implementing those models, we got these results.

Architecture	Accuracy	Precision	Recall	F1 Score
Inception V3	84.00	87.00	84.00	83.00
DenseNet121	90.62	90.03	90.03	90.62
ResNet101_v2	92.00	92.00	92.00	92.00
VGG16	94.00	93.00	96.00	95.00
Proposed_CNN_DenseNet	99.11	98.89	99.01	98.79

Table 3: Accuracy, precision, recall, and F1 score corresponding to different architectures.

Several indicators were used to assess the performance of different neural network topologies. 84.00% accuracy, 87.00% precision, 84.00% recall, and 83.00% F1 score were attained using the Inception V3 model. With an accuracy of 90.62%, precision of 90.03%, recall of 90.03%, and F1 score of 90.62%, the DenseNet121 model outperformed the others. Further development was demonstrated by the ResNet101_v2 model, which scored 92.00% on all criteria. With an accuracy of 94.00%, precision of 93.00%, recall of 96.00%, and F1 score of 95.00%, the Proposed_CNN_Vgg16 model surpassed these. The Proposed_CNN_DenseNet model had the best accuracy, coming in at 99.11%; this model has a precision of 98.89, a recall of 99.01, and an F1 score of 98.79. The suggested models significantly improved overall. All things considered, the suggested models outperformed the existing designs significantly, especially in terms of accuracy.

7.1. Result in Terms of Testing Accuracy and Testing Loss:

Google Colab is used to implement the suggested tasks in Python. First, we used the CNN model to train this dataset. But the outcome was not better. Later, we raised the number of epochs to 30 and made modifications to a few layers to improve the findings. To reflect better outcomes, we are concerned with the loss function, optimizer, epochs, batch size, and layers.

Each experiment was carried out five times in order to test and assess the suggested network's performance. Throughout the training, parameters and hyperparameters were adjusted. The training accuracy and training loss curves acquired after 10 epochs of model training are displayed. Every model has a training accuracy greater than 99% and a training loss of less than 0.03. All the other models, except the exception, had training loss curves and accuracy that were comparable.

	Accuracy	**Loss**
Training	99.27%	0.02
Validation	100%	0.03
Testing	99.11%	0.13

Table 4: Model Evaluation

The machine learning model's performance metrics during its three phases of training, validation, and testing are shown in the table.

99.5% of the training data is correctly classified by the model during the training phase. A loss of 0.015, which is incredibly low, shows how well the model has absorbed the training set.

The validation stage's accuracy of 96.88% shows how well the model performs on the validation set, which is used to fine-tune the model. Although 96.88% accuracy is high, it is less than the training accuracy, suggesting a minor decline in performance and we got a loss of 0.003 which means it performed well.
The model's performance on the test set is represented by its accuracy of 99.11% in the testing state, which is used to evaluate the final model. Furthermore, impressive is the accuracy of 99.11%, which shows how effectively the model generalizes to fresh, untested data. An excellent sign of generalization is the test set loss of 0.027, which is larger than the training loss but lower than the validation loss.

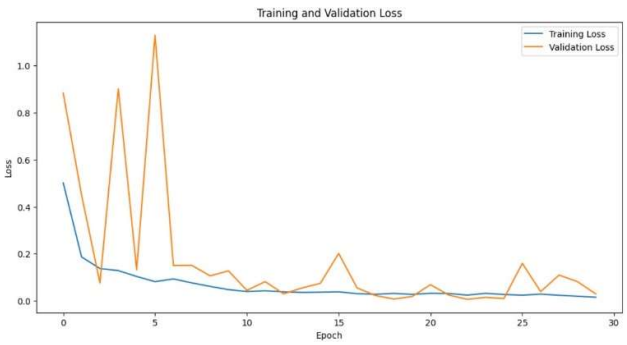

Figure 11: *Training and validation loss*

Figure 12: Training and Testing Accuracy

The above graphs show the accuracy and loss during training and validation for a machine-learning model over a number of epochs.

Within the first epoch, the training loss rapidly drops from a relatively high starting point to almost zero. This sharp decline suggests that the model picks up the training set fast, resulting in little loss during the initial stages of training.
Throughout the training process, the validation loss is relatively flat and begins at a low value. This flat line indicates that while training goes on, the model's performance on the validation set remains mostly unchanged.
Epoch 3 is chosen as the best epoch, signifying the moment at which, in terms of a selected metric (likely validation loss or accuracy), the validation performance is at its best.

7.2. Performance Analysis:

We applied for the CNN model and got the following confusion matrix. Following the conclusion of the training phase, each model was evaluated using the test dataset. F1, the area under the curve, accuracy, recall, and precision were used to validate their performance. The discussion of each performance statistic utilized in this article follows.

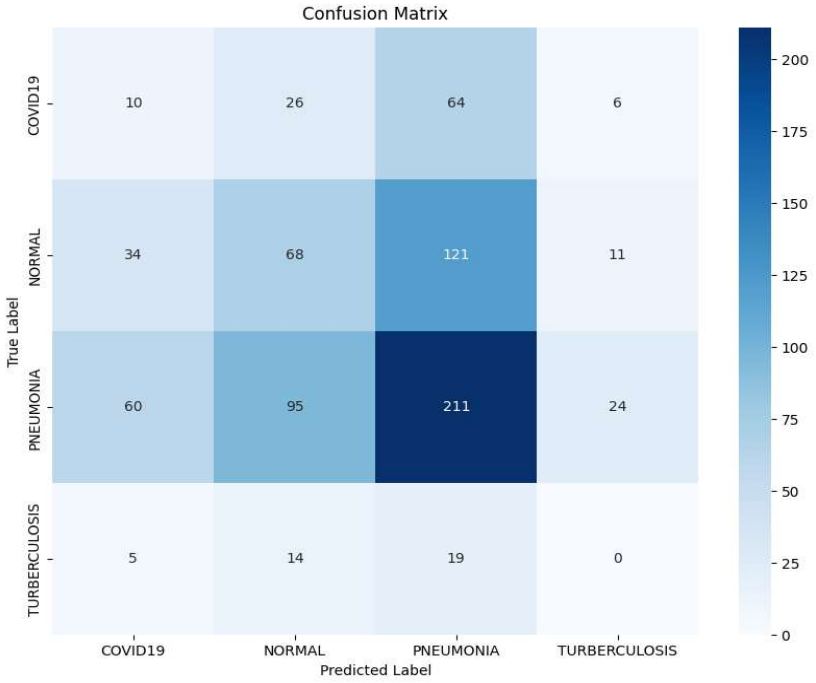

Figure 13: Test Data (CNN Model)

True positive (TP) indicates the number of pneumonia images identified as pneumonia, true negative (TN) indicates the number of normal images identified as normal (healthy), false positive (FP) indicates the number of normal images incorrectly identified as pneumonia images, and false negative (FN) indicates the number of pneumonia images incorrectly identified as normal. These definitions and equations are used to classify patients into healthy and pneumonia patients.

- Accuracy: It indicates the degree to which the measured value resembles a known value.

  Accuracy = (TP + FN)/ (TP + TF + FP + FN) **(2)**
- Precision: This indicates the model's level of accuracy concerning positively anticipated data.

  Precision = TP / (TP + FP) **(3)**
- Recall: After classifying something as positive, it determines how many real positives the model caught (true positives).

  Recall = TP / (TP + FN) **(4)**
- F1: It provides a recall and accurate balance.

  F1 = 2 × (Recall × Precision) / (Recall + Precision) **(5)**

	precision	**recall**	**f1-score**	**Support**
Covid19	0.98	0.96	0.97	106
Normal	0.98	0.52	0.68	234
Pneumonia	0.77	0.99	0.87	390
Tuberculosis	0.93	0.97	0.95	38

Table 5: Classification Report

According to the classification report, we can observe that:

1) With 106 cases, "Covid19" has an f1-score of 0.97, a recall of 0.96, and a precision of 0.98.
2) With 234 instances, "normal" has an f1-score of 0.68, a recall of 0.52, and a precision of 0.98.
3) With 390 instances, "pneumonia" has an f1-score of 0.87, a precision of 0.77, and a recall of 0.99.
4) With 38 instances, "tuberculosis" has an f1-score of 0.95, a recall of 0.97, and a precision of 0.93.

The model performs well in every category, according to this classification report.

7.3. ROC CURVE:

The suggested classifier and the ROC curves for distinct designs are displayed in Figure 14.

The suggested classifier was able to obtain the highest AUC score of 97.7%. Every model displayed a comparable AUC/ROC curve. The weighted classifier produced the greatest results, according to the analysis, with an AUC score of 97.7%, an F1 score of 1.00, and a test accuracy of 98%.

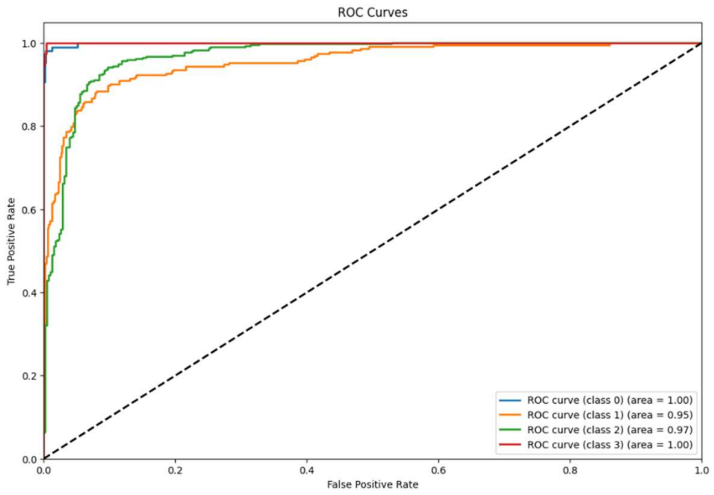

Figure 14: ROC Curve

	precision	recall	f1-score	Support
Covid19	0.98	0.96	0.97	106
Normal	0.98	0.52	0.68	234
Pneumonia	0.77	0.99	0.87	390
Tuberculosis	0.93	0.97	0.95	38

Table 6: Accuracy, precision, recall, F1 score, and AUC score corresponding to architectures.

Classification performance metrics for several classes in a multi-class classification problem are shown in the above table. The classes are Covid19, Normal, Pneumonia, and Tuberculosis.

- Covid19: Excellent recall and precision, but marginally worse recall than other classes, which results in a marginally lower F1-score.
- Normal: Reliable performance with a slightly higher recall than precision, showing high precision and recall.
- Pneumonia: Pneumonia occurrences are flawlessly classified by the model, as seen by perfect scores for all criteria.
- Tuberculosis: Reliable performance is indicated by a balanced and high F1-score, together with excellent precision and recall.

7.4. COMPARATIVE ANALYSIS OF VARIOUS EXISTING METHODS:

The accuracy of the suggested methodology was compared to other current methods. The authors have published all the findings in this section in their individual research.

Model	No. of Images	Precision	Recall	Accuracy	AUC
P Bertin et al. [37]	5232	90.1	93.2	92.8	99
S Candemir et al. [38]	5856	97.0	99.5	96.2	99
B Ergen et al. [39]	5849	96.88	96.83	96.84	96.8
DBS Santos et al. [40]	5840	94.3	94.5	94.4	94.5
Halil Murat Unver et al. [41]	5856	91.3	89.1	84.5	87
Saad Kashem et al. [42]	5247	97	99	98	98
Sanjay Kumar Singh et al. [43]	5232	93.28	**99.6**	96.39	**99.34**
Proposed Methodology	7153	**98**	99	99	99

Table 7: *Comparison of the proposed methodology with different existing methods.*

A DenseNet-121-based model was employed by P Bertin et al. [37]. They stated that their AUC was 98.4%. Regretfully, the paper did not report on the other measures. S Candemir et al. [38] reported a test accuracy of 96.2% using customized CNNs to identify pneumonia. B. Ergen et al. [39] classified pneumonia with 96.84% accuracy by combining features from several deep-learning models. Sanjay Kumar Singh et al. [43] used majority voting to aggregate the outputs of several neural networks to arrive at the final forecast. They received a 99.34 AUC score. Using deep learning-based techniques, DBS Santos et al. [40], Halil Murat Unver et al. [41], and Saad Kashem et al. [42] obtained accuracy of 94.4%, 84.5%, and 98.0%, respectively. The size of the datasets used in each of these studies was comparable. Table 7 summarizes all the results that were previously addressed.

8. DISCUSSIONS:

The Python tasks are implemented using Google Colab. First, we trained this dataset using the CNN model. However, the result was not superior. To enhance the results later, we increased the number of epochs to 30 and changed a few layers. We are concerned with the optimizer, epochs, batch size, layers, and loss function to reflect better results. The performance of the proposed network was tested and evaluated five times for each experiment. Parameters and hyperparameters were changed throughout training. After 10 epochs of model training, the training accuracy and training loss curves are shown. Each model has a training loss of less than 0.03 and a training accuracy of more than 99%. With one notable exception, every other model exhibited equal accuracy and training loss curves. The table displays the performance metrics of the machine learning model throughout each of its three training, validation, and testing phases. During the training phase, the model successfully classifies 99.5% of the training data. The remarkably low loss of 0.015 indicates how well the model has assimilated the training material.

The model's performance on the validation set, which is used to refine the model, is demonstrated by the validation stage's accuracy of 96.88%. We obtained a loss of 0.003, which indicates that it functioned well, and even though 96.88% accuracy is great, it is lower than the training accuracy, indicating a slight reduction in performance.

9. CONCLUSION AND FUTURE SCOPE:

In this paper, we attempt to find a simpler approach for pneumonia detection for chest X-ray images by comparing the different architectures on the same dataset. Based on our findings, we selected the most perfect model, which is easy to train and has one of the best performance metrics. There are several approaches to detecting pneumonia using CNN, and it gives better accuracy, recall, and F1 score. We are using deep learning algorithms that have proven to be more realistic. It can also be observed that the accuracy of the network can be increased by preprocessing techniques. CNN model to provide efficient accuracy and the best solution for pneumonia detection based on chest X-ray images.

For the first, we got less validation accuracy and less training accuracy, so we trained this model repeatedly. After training the CNN model, we got a validation accuracy of 97% and a training accuracy of 99%.

We also show that the simple extension of our algorithm that can perform on multiple diseases outperforms the previous state of the art on Chest X-ray images. The issue of not having enough training for the dataset has also been addressed using data augmentation and transfer learning. In the future, an algorithm that can determine the areas of the lung damaged by pneumonia will be developed. The models of these convolutional neural networks were successfully created using a variety of parameter tuning techniques, including adding dropout, adjusting learning rates, modifying batch sizes, and epoch counts, adding more complex fully connected layers, and adjusting different stochastic gradient optimizers [36].

REFERENCES:

[1] Marrie, T. (1994). Community-acquired pneumonia. Clinical Infectious Diseases. 18(4), 501-513.

[2] CDC, 2017. URL https://www.cdc.gov/features/ pneumonia/index.html.

[3] Sayed E., et. al, "Computer-aided Diagnosis of Human Brain Tumor through MRI: A Survey and a new algorithm. Expert System with Applications (41): 2014.

[4] C. Ciresan, U. Meier, J. Masci, L.M. Gambardella, J. Schmidhuber, High-performance neural networks for visual object classification, (2011), arXiv preprint arXiv:1102.0183.

[5] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. 2014. CNN features off-the-shelf: an astounding baseline for recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 806813.

[6] Nijhawan, Rahul, Rose Verma, Shashank Bhushan, Rajat Dua, and Ankush Mittal." An Integrated Deep Learning Framework Approach for Nail Disease Identification." In Signal-Image Technology and Internet Based Systems (SITIS), 20

[7] Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8): 500–510. doi:10.1038/s41568-018-0016-5

[8] Gang, Peng, Wang Zhen, Wei Zeng, Yuri Gordienko, Yuriy Kochura, Oleg Alienin, Oleksandr Rokovyi, and Sergii Stirenko. "Dimensionality reduction in deep learning for chest x-ray analysis of lung cancer." In 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), pp. 878-883. IEEE, 2018.

[9] Khobragade, Shubhangi, Aditya Tiwari, C. Y. Patil, and Vikram Narke. "Automatic detection of major lung diseases using Chest Radiographs and classification by a feed-forward artificial neural network." In 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), pp. 1-5. IEEE, 2016.

[10] Paras Lakhani and Baskaran Sundaram. 2017. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284, 2 (2017), 574582.

[11] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, and others. 2017. Chexnet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017).

[12] Udeshani, K. A. G., R. G. N. Meegama, and T. G. I. Fernando. "Statistical feature-based neural network approach for the detection of lung cancer in chest x-ray images." International Journal of Image Processing (IJIP) 5, no. 4 (2011): 425-434.

[13] Guan, Qingji, Yaping Huang, Zhun Zhong, Zhedong Zheng, Liang Zheng, and Yi Yang. "Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification." arXiv preprint arXiv:1801.09927 (2018).

[14] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, IEEE Trans. Med. Imaging 35 (5) (2016) 1207–1216.

[15] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in Advances in neural information processing systems, (2012), pp. 1097–1105.

[16] Benjamin Antin, Joshua Kravitz, and Emil Martayan. 2017. Detecting Pneumonia in Chest X-Rays with Supervised Learning. (2017).

[17] Douarre, C.; Schielein, R.; Frindel, C.; Gerth, S.; Rousseau, D. Transfer learning from synthetic data applied to soil–root segmentation in x-ray tomography images. J. Imaging 2018, 4, 65. [CrossRef]

[18] Zhang, Y.; Wang, G.; Li, M.; Han, S. Automated classification analysis of geological structures based on images data and deep learning model. Appl. Sci. 2018, 8, 2493. [CrossRef]

[19] Wang, Y.; Wang, C.; Zhang, H. Ship classification in high-resolution SAR images using deep learning of small datasets. Sensors 2018, 18, 2929. [CrossRef] [PubMed]

[20] Sun, C.; Yang, Y.; Wen, C.; Xie, K.; Wen, F. Voiceprint identification for a limited dataset using the deep migration hybrid model based on transfer learning. Sensors 2018, 18, 2399. [CrossRef] [PubMed]

[21] Chen, Z.; Zhang, Y.; Ouyang, C.; Zhang, F.; Ma, J. Automated landslides detection for mountain cities using multi-temporal remote sensing imagery. Sensors 2018, 18, 821. [CrossRef]

[22] Razzak, M.I.; Naz, S.; Zaib, A. Deep learning for medical image processing: Overview, challenges, and the future. In Classification in BioApps; Springer: Cham, Switzerland, 2018. pp. 323–350.

[23] Roth HR, et al. A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations. International conference on medical image computing and computer-assisted intervention; Springer, Cham; 2014.

[24] Rajaraman S, Candemir S, Kim I, et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. Appl Sci. 2018;8(10):1715.

[25] Abiyev, R.H.; Ma'aitah, M.K.S. Deep convolutional neural networks for chest diseases detection. J. Healthc. Eng. 2018, 2018, 4168538. [CrossRef] [PubMed]

[26] Stephen, O.; Sain, M.; Maduh, U.J.; Jeong, D.U. An efficient deep learning approach to pneumonia classification in healthcare. J. Healthc. Eng. 2019, 2019, 4180949. [CrossRef] [PubMed]

[27] Shah, S., Mehta, H. and Sonawane, P., 2020, August. Pneumonia detection using convolutional neural networks. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 933-939). IEEE.

[28] Račić, L., Popović, T. and Šandi, S., 2021, February. Pneumonia detection using deep learning based on convolutional neural network. In *2021 25th International Conference on Information Technology (IT)* (pp. 1-4). IEEE.

[29] Xu, Y., Jia, Z., Ai, Y., Zhang, F., Lai, M., Eric, I. and Chang, C., 2015, April. Deep convolutional activation features for large-scale brain tumor histopathology image classification and segmentation. In 2015 IEEE international conference on acoustics, speech, and signal processing (ICASSP) (pp. 947-951). IEEE.

[30] Rubin, J., Sanghavi, D., Zhao, C., Lee, K., Qadir, A. and Xu-Wilson, M., 2018. Large scale automated reading of frontal and lateral chest X-rays using dual convolutional neural networks. *arXiv preprint arXiv:1804.07839*.

[31] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[32] Xu, Y., Jia, Z., Ai, Y., Zhang, F., Lai, M., Eric, I. and Chang, C., 2015, April. Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In *2015 IEEE international conference on acoustics, speech, and Signal Processing (ICASSP)* (pp. 947-951). IEEE.

[33] Wang, S.H., Xie, S., Chen, X., Guttery, D.S., Tang, C., Sun, J., and Zhang, Y.D., 2019. Alcoholism identification based on an AlexNet transfer learning model. *Frontiers in psychiatry*, *10*, p.454348.

[34] Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, V2, doi: 10.17632/rscbjbr9sj.2

[35] Rahman, T., Chowdhury, M.E., Khandakar, A., Islam, K.R., Islam, K.F., Mahbub, Z.B., Kadir, M.A. and Kashem, S., 2020. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Applied Sciences*, *10*(9), p.3233.

[36] Du, S. S., Zhai, X., Poczos, B., Singh, A.: Gradient Descent Provably Optimizes OverParameterized Neural Networks (2018). arXiv preprint arXiv:1810.02054

[37] Cohen, J.P.; Bertin, P.; Frappier, V. Chester: A Web Delivered Locally Computed Chest X-Ray Disease Prediction System. arXiv 2019, arXiv:1901.11210

[38] Rajaraman, S.; Candemir, S.; Kim, I.; Thoma, G.; Antani, S. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. Appl. Sci. 2018, 8, 1715.

[39] To˘gaçar, M.; Ergen, B.; Cömert, Z. A deep feature learning model for pneumonia detection applying a combination of mRMR feature selection and machine learning models. IRBM 2019.

[40] Saraiva, A.; Santos, D.; Costa, N.J.C.; Sousa, J.V.M.; Ferreira, N.F.; Valente, A.; Soares, S. Models of Learning to Classify X-ray Images for the Detection of Pneumonia using Neural Networks. 2019.

[41] Ayan, E.; Ünver, H.M. Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning. In Proceedings of the 2019 Scientific Meeting on Electrical-Electronics &

Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 2–26 April 2019. pp. 1–5.

[42] Rahman, T.; Chowdhury, M.E.; Khandakar, A.; Islam, K.R.; Islam, K.F.; Mahbub, Z.B.; Kadir, M.A.; Kashem, S. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray. Appl. Sci. 2020, 10, 3233

[43] Chouhan, V.; Singh, S.K.; Khamparia, A.; Gupta, D.; Tiwari, P.; Moreira, C.; Damaševiˇcius, R.; de Albuquerque, V.H.C. A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images. Appl. Sci. 2020, 10, 559.

[44] Dataset link: https://www.kaggle.com/datasets/jtiptj/chest-xray-pneumoniacovid19tuberculosis.

[45] Hashmi, Mohammad Farukh, Satyarth Katiyar, Avinash G. Keskar, Neeraj Dhanraj Bokde, and Zong Woo Geem. "Efficient pneumonia detection in chest Xray images using deep transfer learning." Diagnostics 10, no. 6 (2020): 417.

[46] Hashmi, Mohammad Farukh, Satyarth Katiyar, Abdul Wahab Hashmi, and Avinash G. Keskar. "Pneumonia detection in chest X-ray images using compound scaled deep learning model." AUTOMATIKA 62, no. 3-4 (2021): 397-406.