Detection of Pneumonia and Tuberculosis using Deep Learning approach

Alekhya Ayinam
Computer Science and Engineering
PVP Siddhartha Institute of Technology
Vijayawada, India
alekhya200127@gmail.com

Gudimetla Siri Sai Sri Ram
Computer Science and Engineering
PVP Siddhartha Institute of Technology
Vijayawada, India
19501a0563@pvpsiddhartha.ac.in

Gali Bhavana
Computer Science and Engineering
PVP Siddhartha Institute of Technology
Vijayawada, India
19501a0554@ pvpsiddhartha.ac.in

Gandamala Mikihil
Computer Science and Engineering
PVP Siddhartha Institute of Technology
Vijayawada, India
19501a0555@ pvpsiddhartha.ac.in

Aditya Gurram
Computer Science and Engineering
PVP Siddhartha Institute of Technology
Vijayawada, India
19501a0502@ pvpsiddhartha.ac.in

S. Madhavi
Computer Science and Engineering
PVP Siddhartha Institute of Technology
Vijayawada, India
mmadhavi@pvpsiddhartha.ac.in

***Abstract*—** Pneumonia and tuberculosis are two major respiratory diseases that can be diagnosed through chest X-rays. However, accurate and efficient detection of these diseases is a challenge for radiologists, as it requires careful examination of multiple images. A deep learning approach can help to automate this process and improve the accuracy of disease detection. The aim of this study is to develop a machine-learning model that can detect and classify cases of pneumonia and tuberculosis using chest X-rays. Specifically, aim to investigate the effectiveness of transfer learning and the ResNet50 model for this task. A publicly available dataset of chest X-rays from Kaggle is used to train and test the model. The dataset includes images from patients with and without pneumonia and tuberculosis. Transfer learning with the ResNet50 model a multilayer CNN model along with a residual block is used, which has been shown to be effective for image classification tasks. Trained the model using a combination of data augmentation and fine-tuning techniques. For quantitative treatment, accuracy is used as the main metric for evaluating the model's performance, and also used precision, recall, and F1 score to assess the model's ability to detect cases of pneumonia and tuberculosis. The model achieved an accuracy of 96.59% for detecting cases of pneumonia tuberculosis. Additionally, the precision, recall, and F1 score for both diseases were above 0.90. These results demonstrate the effectiveness of transfer learning and the ResNet50 model for disease detection in chest X-rays. Furthermore, the model outperformed other state-of-the-art models on this dataset.

***Keywords*—** *Deep learning, ResNet50, Convolutional Neural Network (CNN), chest X-rays, residual blocks, Transfer learning, precision, recall, F1score, pneumonia, tuberculosis*

I. INTRODUCTION

Deep learning models can solve this problem of detecting pneumonia and tuberculosis through chest X-rays by giving precise results as numerous machine learning models failed to detect diseases through chest X-rays [13]. A literature survey showed that Deep learning models are way better than ML models in identifying a disease by classifying the images [11] [14]. A deep learning model like densely connected CNN can be used for best feature extraction [1]. Transfer learning is a process of pre-training an existing model for accuracy and efficiency [4]. Transfer learning helps in many ways like taking less time for training and giving the best results with fewer data.

A transfer learning technique is used to get efficient results [13]. A pre-trained architecture Resnet50 is used to extract features and classify the disease. This Neural Network consists of 50 layers including the pooling layers. The convolutional layers can classify images automatically by extracting features [15]. Rectified Linear Unit (ReLu) is an activation function to upgrade the non-linearity.

II. DATASET

The chest X-ray dataset used in the study contains a total of 6,560 images, with three classes: Pneumonia, Normal, and Tuberculosis. Specifically, there are 4,285 images of Pneumonia, 1,575 images of Normal, and 700 images of Tuberculosis. To split the dataset for training and testing, a 75/25 split is used, with 75% of the data (4,920 images) used for training and 25% of the data (1,640 images) used for testing. The distribution of the three classes in the dataset is shown in Fig. 1, with the x-axis indicating the disease and the y-axis indicating the number of images for each respective disease. Overall, this dataset is publicly available on Kaggle and is commonly used for developing and evaluating machine learning models for chest x-ray image classification.

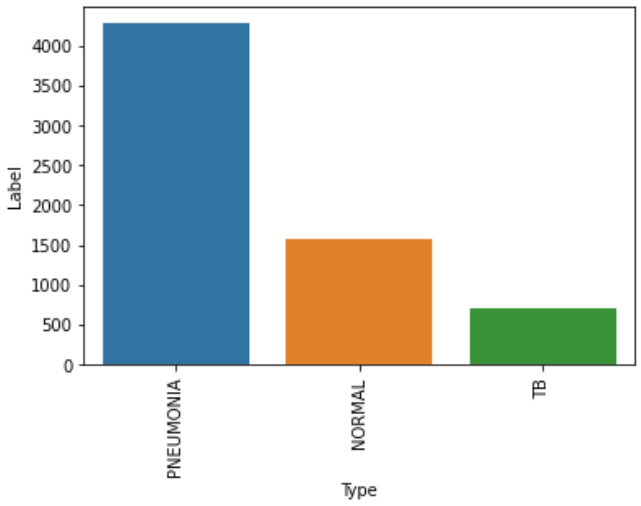

Fig. 1. Graph showing a number of images for the respective class.

A. Dataset splitting

In this, a method to divide the data into training and testing called train_test_split() in sklearn is used.

train,test=train_test_split(data,test_size=0.25,random_state=42)

The hyperparameter test_size=0.25 means 25% of the dataset is used for testing and the other 75% is for training. randome_state=42 for consistent results.

A. Pre-processing

For image pre-processing. ImageDataGenerator which is imported from TensorFlow is used. It consists of a built-in function called preprocessing_function. The input image is passed through a preprocessing function which has a set of functions like rotation, shifting the pixels, flipping along the axis, etc to perform data augmentation automatically [2].

Another function from ImageDataGenerator class called flow_from_dataframe is used for generating the batches of augmented data and reading images directly from the directory [13].

The hyperparameter in this function used is target_size to resize the image into 227X227 to make it fit the model [8].

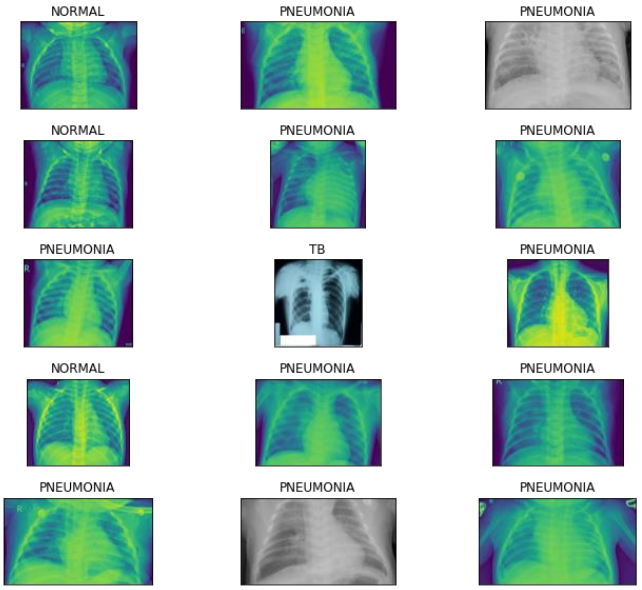

Fig. 2. Sample images from the dataset

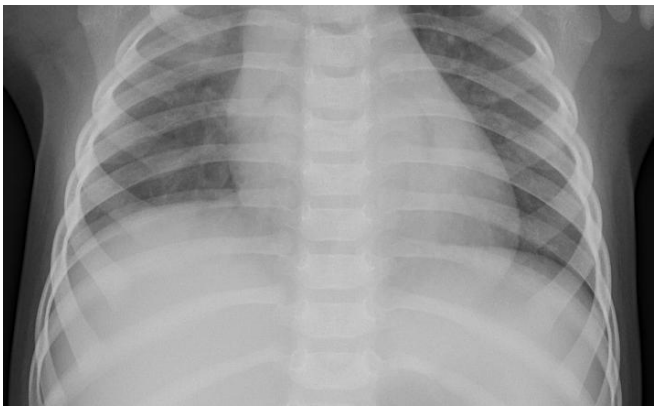

Fig. 3. 712 X 439 size chest x-ray

Fig. 4. 227 X 227 size chest x-ray

Fig. 4 is a resized image of Fig. 3 by pre-processing.

III. MODEL

A. Why ResNet-50 ?

Many different architectures were used before for classification but they were not efficient and the results were not up to the mark [10]. For instance, the feature extractions using VGG16 do not provide desired features to classify the chest X-ray images as it has a smaller network [12]. Another example is DenseNet- 121 which gives similar results to VGG16 but it takes a long time to train. The resNet-50 model can be used to overcome these problems.

ResNet stands for Residual Network which is a type of convolutional neural network(CNN) [3]. ResNet-50 is a 50 layers network consisting of skip functions to skip the layers. This function helps to prevent the Vanishing Gradient problem, this problem arises when there the network is huge i.e., with multiple layers when a backpropagation is performed the gradient gets decreased which results in low performance [7]. The residual blocks in Fig. 5 contain skip functions that connect the input of one layer to the output of the next layer which helps to skip the layers while backpropagation prevents the problem of decreasing the gradient and providing accurate results. ResNet is already trained on the imagenet dataset so it will be easy to pre-train it for detecting pneumonia from chest X-rays.

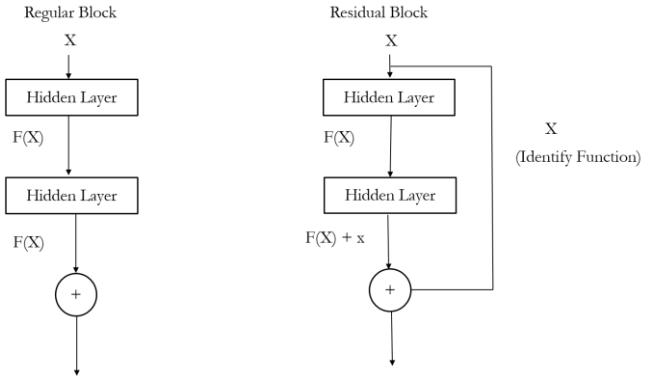

Fig. 5. Regular block vs Residual blocks.

B. ResNet-50 architecture

ResNet-50 architecture Fig. 6 consists of 50 layers. It has bottleneck residual blocks which use 1X1 convolutions, this helps in faster training of each layer. The architecture of ResNet50 can be divided into several stages, each containing multiple residual blocks. Each residual block contains convolutional layers, batch normalization layers, ReLU activation layers, and shortcut connections that skip one or more layers. The shortcut connections enable gradient flow through the network and help to alleviate the vanishing gradient problem.

The ResNet-50 has the following layers:

- The input layer, which takes in an image of size 224x224x3 (width x height x channels).
- A convolutional layer with 64 filters of size 7x7 and a stride of 2, followed by batch normalization and ReLU activation.
- A max pooling layer with a pool size of 3x3 and a stride of 2.
- Four stages, each containing multiple residual blocks. The residual blocks in each stage are 3, 4, 6, and 3, respectively.
- The final stage contains a global average pooling layer followed by a fully connected layer with 1000 units, which represents the number of classes in the ImageNet dataset.
- The output layer uses a softmax activation function to output the predicted class probabilities.
- The sum of the above layers i.e 1+9+12+18+9+1 gives 50 layers.

  Each residual block in ResNet50 contains multiple convolutional layers and shortcut connections, which enable the network to learn complex representations of images and achieve high accuracy on tasks such as image classification.

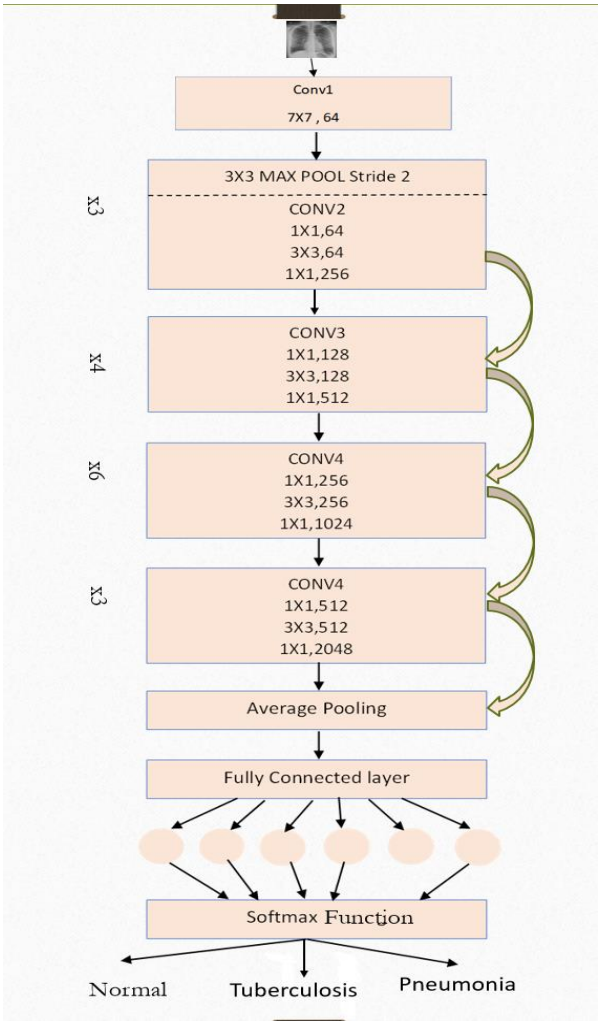

Fig. 6. ResNet-50 architecture

C. Layer Description

After preprocessing/ data augmentation the input image is resized to standard shape 227X227 and it is passed to the first convolutional layer with 3X3,64 size, where feature extraction is done. This layer is used to extract features from the input image and also reduces the size [12]. Here the important patterns in the image are recognized [9]. A mathematical operation in Fig. 7 is performed between an image matrix and a filter. An image matrix of AXBXC and filter size of DXEXF produces an output size of (A-D+1)*(B-E+1)*1.
Where A=height, B=width, C=Depth .

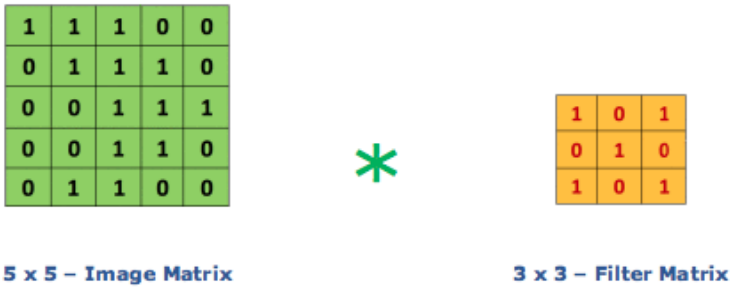

Fig. 7. Convolution operation

Next obtained feature maps are sent to max pooling, this layer with 3X3 window size, the size of the input image reduces and important features are extracted. In this way, the input image passes through all the convolution and max pooling layers to get the important details of the image. These layers are activated using the activation function ReLu. Average pooling is applied before passing it to flatten layer. Stride and padding are applied to convolution layers [12]. Stride is the number of pixels that jump over the input matrix. Padding is used when the kernel does not match the input image. To maintain the original size a method called zero padding is used.

From the average pooling, the output is passed to flatten layer where the image is flattened into a vector. Then from flatten layer, it is passed to the fully connected layer. This layer classifies images by using extracted features obtained from the previous layer. The main purpose of the fully connected layer returns a set of class scores and perform classification based on the image features. At last, a softmax function is used as a classifier. The softmax function is an activation function used to distribute the probabilities among classes. It is a classifier to predict the probabilities of the classes and draws a conclusion based on the probabilities of pneumonia, normal, and tuberculosis.

IV. Methodology

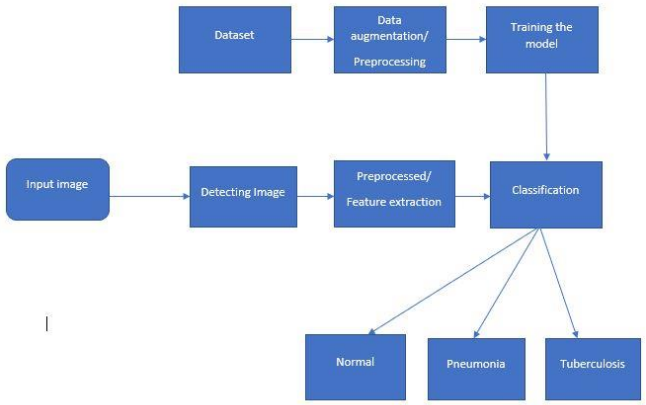

Fig. 8. Process of detecting the presence of disease

The dataset taken from the Kaggle is first preprocessed and augmented. By using this data, the model has been trained here the model is ResNet-50 which is the pre-trained model and it is trained according to the requirements. The model is trained to classify the disease. When the input image is fed into the model the model detects the image and feature extraction is performed based on that the model classifies it as pneumonia, normal, or tuberculosis.

V. Testing

A. Graphs

a) Accuracy Graph(Fig.9)

The accuracy obtained at every epoch is plotted as a graph, with the epochs on the x-axis and the accuracy on the y-axis. The training accuracy is indicated by the blue curve which is increasing and shows that the model learning rate has increased through training data, the orange curve indicates validation accuracy and is calculated on the basis of the generalization of the model.

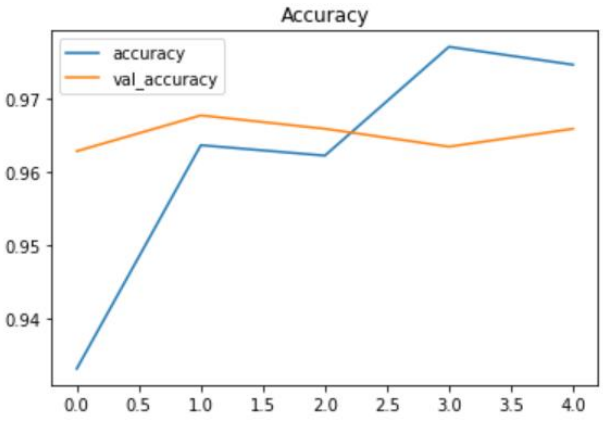

Fig. 9. Accuracy graph

b) Loss Graph(Fig.10)

Loss obtained at every epoch is plotted as a graph the epochs are on the x-axis and loss is shown on the y-axis. Here training loss is indicated by the blue curve which is decreasing it shows the training data is fit to the model and the orange curve indicates validation loss which describes how well the model fits new data.

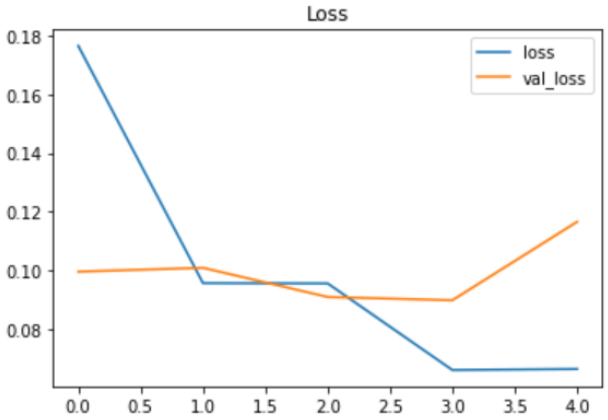

Fig. 10. Loss graph

B. Test cases

Three Parameters i.e. input, actual output, and predicted output are considered to tabulate the below table.

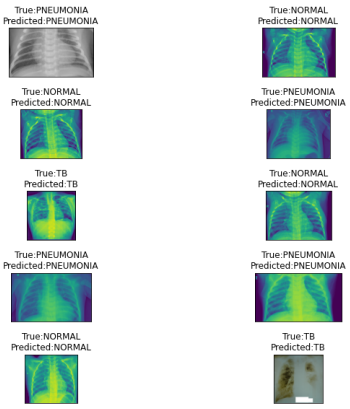

Fig. 11. True and Predicted values

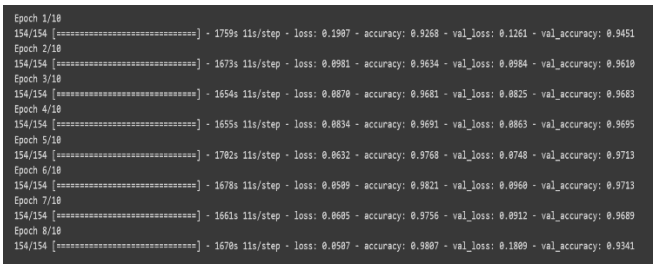

Fig. 12. Training of model with 10 epochs

Here early stopping function is used so the epochs stopped at the 8th epoch as val_accuracy decreased from the 7th epoch Fig. 12.

VI. Results

The Test Accuracy and Test Loss are calculated, where we tested the model with different test cases. Both accuracy and loss resulted in 93.41 and 0.18086 respectively.

Test Loss: 0.18086
Test Accuracy: 93.41

A. *Classification reports*

Classification report Fig. 13 is a performance evaluation metric that is used to show all performance metrics like precision, recall, F1 score, and accuracy in a single report. This report also shows Support for the trained classification model. Besides these metrics, it also shows weighted and macro averages.

	precision	recall	f1-score	support
NORMAL	0.98	0.93	0.95	394
PNEUMONIA	0.98	0.99	0.98	1086
TB	1.00	1.00	1.00	160
accuracy			0.98	1640
macro avg	0.98	0.97	0.98	1640
weighted avg	0.98	0.98	0.98	1640

Fig. 13. Report of classification

VII. Conclusion

The proposed solution to detect pneumonia and tuberculosis based on chest X-rays gives accurate results. The deep learning approach used in this paper gave the best results by extracting important features of the image using the convolution layer and training the model using a pre-trained model ResNet-50. By using early stopping the loss is diminished and increased in accuracy while training the model.

6560 images of chest X-rays are used which consist of 3 classes normal, pneumonia, and tuberculosis. By using this dataset, the model is trained and classification is performed on the basis of training.

The model is trained on 75% of images and 25% for testing. By using accuracy, precision, F1 score the performance of the model is calculated. The experimental results achieved an accuracy of 96.59% for the proposed model ResNet50.

Overall, the proposed deep learning approach shows promise in detecting pneumonia and tuberculosis from chest X-rays, but further research is needed to validate the approach's generalizability and clinical relevance.

References

[1] Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R., & Mittal, A. (2019). Pneumonia detection using CNN based feature extraction. In 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) (pp. 1–7). IEEE.

[2] Li, Y., Zhang, Z., Dai, C., Dong, Q., & Badrigilan, S. (2020). Accuracy of deep learning for automated detection of pneumonia using chest X-ray images: a systematic review and meta-analysis. Computer. Biol. Med., 123, 103898. https://doi.org/10.1016/j.compbiomed.2020.103898

[3] Praveen, S. P., Srinivasu, P. N., Shafi, J., Wozniak, M., & Ijaz, M. F. (2022). ResNet-32 and FastAI for diagnoses of ductal carcinoma from 2D tissue slides. Scientific Reports, 12(1), 20804.

[4] Elshennawy, N. M., & Ibrahim, D. M. (2020). Deep-pneumonia framework using deep learning models based on chest x-ray images. Diagnostics, 10(9), 649. https://doi.org/10.3390/diagnostics10090649

[5] Demner, D., Kohli, M. D., Rosenman, M. B., Shooshan, S. E., Rodriguez, L., Antoni, S., Thomas, G. R., & McDonald, C. J. (2015). Preparing a collection of radiology examinations for distribution and retrieval. J. Am. Med. Inform. Assoc., 23(2), 304–310.

[6] Cherian, T., Mulholland, E. K., Carlin, J. B., Ostensen, H., Amin, R., Campo, M. D., Greenberg, D., Lagos, R., Lucero, M., Madhi, S. A., & O'Brien, K. L. (2005). Standardized interpretation of pediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. Bull. World Health Organ., 83, 353–359.

[7] Srinivasu, P. N., Shafi, J., Krishna, T. B., Sujatha, C. N., Praveen, S. P., & Ijaz, M. F. (2022). Using Recurrent Neural Networks for Predicting Type-2 Diabetes from Genomic and Tabular Data. Diagnostics, 12(12), 3067.

[8] Wang, X., Peng, Y., Lu, L., Lu, Z., & Summers, R. M. (2017). Chestx-ray8: Hospital scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In 2017 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2097–2106).

[9] Swamy, S. R., Praveen, S. P., Ahmed, S., Srinivasu, P. N., & Alhumam, A. (2023). Multi-features disease analysis based smart diagnosis for covid-19. Computer Systems Science and Engineering, 869-886.

[10] Roth, H.R., Lu, L., Seff, A., Cherry, K.M., Hoffman, J., Wang, S., Liu, J., Turkey, E., Summers, R.M. (2014). A new 2.5 d representation for lymph node detection using random ets of deep convolutional neural network observations. In Proceedings of the 17th International Conference on Medical Image Computing and Computer Assisted Intervention (pp. 520-527). Springer.

[11] Praveen, S.P., Jyothi, V.E., Anuradha, C., VenuGopal, K., Shariff, V., & Sindhura, S. (2022). Chronic Kidney Disease Prediction Using ML-Based Neuro-Fuzzy Model. International Journal of Image and Graphics, 23(4), 2340013.

[12] Jung, H., Kim, B., Lee, I., Lee, J., & Kang, J. (2018). Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method. BMC Medical Imaging, 18(1), 1-10.

[13] Jaiswal, A.K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J.J.P.C. (2019). Identifying pneumonia in chest X-rays: A deep learning approach. Measurement, 145, 511-518.

[14] Anuradha, C., Swapna, D., Thati, B., Sree, V.N., & Praveen, S.P. (2022, January). Diagnosing for Liver Disease Prediction in Patients Using Combined Machine Learning Models. In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 889-896). IEEE.

[15] Nasrullah, N., Sang, J., Alam, M.S., Mateen, M., Cai, B., & Hu, H. (2019). Automated lung module detection and classification using deep learning combined with multiple strategies. Sensors, 19(3722), 1-19.