International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 6, September-October 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Role of Advanced Machine Learning Techniques and Deep Learning Approach Based Decision Support System for Accurate Diagnosis of Severe Respiratory Diseases

Patel Smitkumar Hareshbhai

Student, Department of Biomedical Engineering, Government Engineering College, Gandhinagar, Gujarat, India

ABSTRACT

Accurate Medical diagnosis is not always possible at every medical center, especially in the Developing Countries where poor healthcare services and lack of advanced diagnostic methods and equipments affects procedures of medical diagnosis .Also, physician intuition and experience are not always sufficient to achieve high quality medical procedures results. Therefore, diagnostic errors and undesirable results are reasons for a need for Machine Learning Techniques based decision support system, which in turns reduce diagnostic errors, increasing the patient safety and save lives. This research focuses on this aspect of Medical diagnosis by learning pattern through the collected dataset of respiratory diseases such as: pneumonia and Covid-19, also consist implementation and test of intelligent medical decision support system to assist physicians and radiologists can deliver great assistance by improving their decision-making ability. In this Research paper, the proposed System use Neural network Resnet-50 and transfer learning technique to classify these severe diseases and performs evaluation of precision, accuracy, specificity of Decision support system.

KEY WORDS: *Medical diagnosis, Machine learning, Respiratory Disease, Decision Support System, Transfer Learning, Neural Network*

How to cite this paper: Patel Smitkumar Hareshbhai "Role of Advanced Machine Learning Techniques and Deep Learning Approach Based Decision Support System for Accurate Diagnosis of Severe Respiratory Diseases" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-6, October 2021, pp.1545-1552, URL: www.ijtsrd.com/papers/ijtsrd47655.pdf

IJTSRD47655

1. INTRODUCTION

The major challenge facing the healthcare industry is the provision for quality services at affordable costs. A quality implies diagnosing patients correctly and treating them effectively. Poor clinical decisions can lead to disastrous results which is unacceptable. Medical diagnosis can be performed in an automatic way with the use of computer-based systems or algorithms. Such systems are usually called diagnostic decision support systems (DDSSs) or medical diagnosis systems (MDSs), which fall under the more general category of decision support system (DSS). The aim of these types of systems is to guide the physicians through the systematic differential diagnosis process.

Respiratory Diseases are very serious health problems in the life of people. These diseases include chronic obstructive pulmonary disease, pneumonia, pleural effusion, COVID-19 and other lung diseases. The timely diagnosis of various Respiratory diseases is very important. There are different types of Lung diseases which can be categorized by Accurate analysis of X-rays. Chest radiography is an economical and easy- to-use medical imaging and diagnostic technique. The technique is the most Innovative that used diagnostic tool in medical practice and has an important role in the diagnosis of the lung disease. Many image processing , Machine Learning and Deep Learning Approaches have been developed for this purpose.

Deep learning-based techniques has been considered to be one of the most emerging & advanced techniques in recent years in computer aided systems. This allows Operators to visualize the results immediately and with great accuracy. Additionally, intuition and experience are frequently used by physicians to reach the medical verdicts instead of knowledge base. Consequently, unsolicited preconceptions and excessive medical cost are produced which may portend the patient's lives. In this paper, It is proposed to build a Neural Network based decision support system for heart diseases diagnosis. Neural Network is considered to be extremely Accurate Systems in image processing as it provides accurate assessment of a disease by both image acquisitions and image interpretation.

I. Machine learning approach

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people.

Although machine learning is a field within computer science, it differs from traditional computational approaches.

Machine learning algorithms instead allow for computers to train on data inputs and use statistical analysis in order to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on data inputs. Machine learning has become a powerful tool for analyzing medical domains, assessing the importance of clinical parameters, and extracting medical knowledge for outcomes research. ML approaches that result in good overall performance and provide medical staff with interpretable prognostic information, providing the ability to support decisions and to reduce the number of medical tests for a reliable prognosis are also desirable. A measure of reliability of the diagnosis or prognosis is also important, because this would give medical staff sufficient confidence to put the new approach into practice.

II. Introduction To Deep Learning

Deep learning attempts to imitate how the human brain can process light and sound stimuli into vision and hearing. A deep learning architecture is inspired by biological neural networks and consists of multiple layers in an artificial neural network made up of hardware and GPUs.

Deep learning attempts to imitate how the human brain can process light and sound stimuli into vision and hearing. A deep learning architecture is inspired by biological neural networks and consists of multiple layers in an artificial neural network made up of hardware and GPUs.

Among the machine learning algorithms that are currently being used and developed, deep learning absorbs the most data and has been able to beat humans in some cognitive tasks. Because of these attributes, deep learning has become the approach with significant potential in the artificial intelligence. Computer vision and speech recognition have both realized significant advances from deep learning approaches.

III. Neural Network

Deep learning is a machine learning method inspired by the deep structure of a mammal brain. The deep structures are characterized by multiple hidden layers allowing the abstraction of the different levels of the features.

This learning algorithm is seen as an unsupervised single layer greedily training where a deep network is trained layer by layer. Because this method became more effective, it has been started to be used for training many deep networks. One of the most powerful deep networks is the convolutional neural network that can include multiple hidden layers performing convolution and subsampling in order to extract low to high levels of features of the input data. This network has shown a great efficiency in different areas, particularly, in computer vision, biological computation, fingerprint enhancement, and so on. Basically, this type of networks consists of three layers: convolution layers, subsampling or pooling layers, and full connection layers. It relies on training data to learn and improve their accuracy over time. This may use for predictive modeling, adaptive control where they can be trained via dataset.

2. METHODOLOGY

I. Supervised Learning

In supervised learning, the computer is provided with example inputs that are labelled with their desired outputs. The purpose of this method is for the algorithm to be able to "learn" by comparing its actual output with the "taught" outputs to find errors, and modify the model accordingly. Supervised learning therefore uses patterns to predict label values on additional unlabeled data.

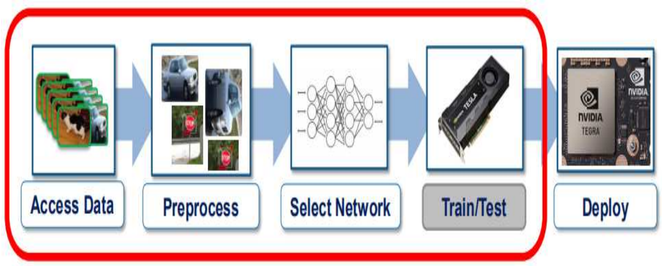


Figure 1. Deep Learning Workflow

II. Supervised Learning Workflow

In supervised learning, the computer is provided with example inputs that are labelled with their desired outputs. The purpose of this method is for the algorithm to be able to "learn" by comparing its actual output with the "taught" outputs to find errors, and modify the model accordingly. Supervised learning therefore uses patterns to predict label values on additional unlabeled data.

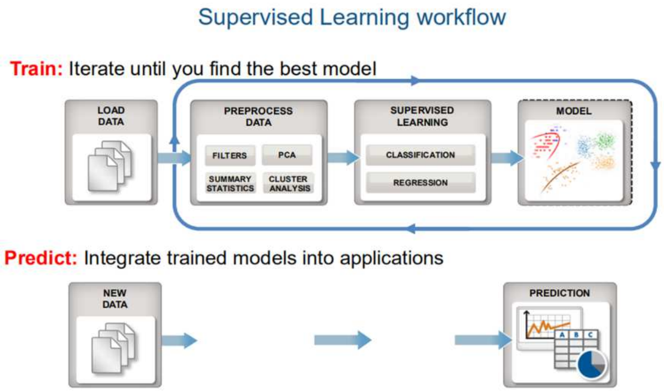


Figure 2. Supervised Learning Workflow

III. Transfer Learning in Lung Diseases Classification

Transfer learning is a very popular approach in computer vision related tasks using deep neural networks when data resources are scarce. Therefore, to launch a new task, we incorporate the pre-trained models skilled in solving similar problems. This method is crucial in medical image processing due to the shortage of real samples.

In deep neural networks, feature extraction is carried out but passing raw data through models specialized in other tasks. Here, we can refer to deep learning models such as ResNet-50, where the last layer information serves as input to a new classifier. Transfer learning in deep learning problems can be performed using a common approach called pre-trained model's approach. Restructured Model states that pre-trained model can produce a starting point for another model used in a different task. This involves incorporation of the whole model or its parts.

The adopted model may or may not need to be refined on the input-output data for the new task. The third option considers selecting one of the available models. It is very common that research institutions publish their algorithms trained on challenging datasets which may fully or partially cover the problem stated by a new task.

IV. ResNet-50 Neural Network

ResNet, short for Residual Network is a specific type of neural network that was introduced in 2015. ResNet-50 is a convolutional neural network that is 50 layers deep.

ResNet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8 x 109 Floating points operations. It is a widely used ResNet model and we have explored ResNet50 architecture in depth.

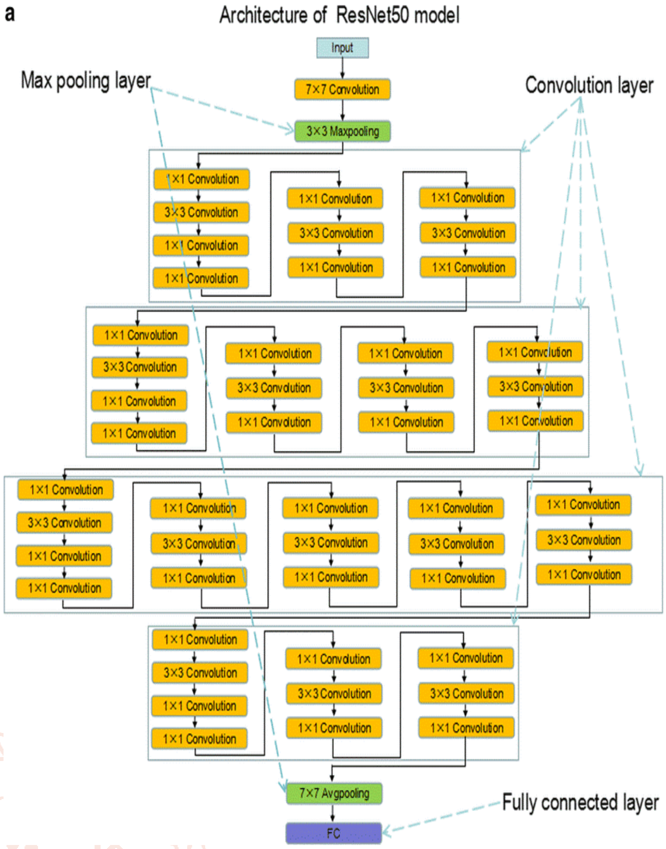


Figure 3. ResNet-50 Architecture

The architecture of ResNet50 has 4 stages as shown in the diagram below. The network can take the input image having height, width as multiples of 32 and 3 as channel

width. For the sake of explanation, we will consider the input size as 224 x 224 x 3. Every ResNet architecture performs the initial convolution and max-pooling using 7×7 and 3×3 kernel sizes respectively. Afterward, Stage 1 of the network starts and it has 3 Residual blocks containing 3 layers each. The size of kernels used to perform the convolution operation in all 3 layers of the block of stage 1 are 64, 64 and 128 respectively. The curved arrows refer to the identity connection. The dashed connected arrow represents that the convolution operation in the Residual Block is performed with stride 2, hence, the size of input will be reduced to half in terms of height and width but the channel width will be doubled. As we progress from one stage to another, the channel width is doubled and the size of the input is reduced to half.

For each residual function F, 3 layers are stacked one over the other. The three layers are 1×1, 3×3, 1×1 convolution. The 1×1 convolution layers are responsible for reducing and then restoring the dimensions. The 3×3 layer is left as a bottleneck with smaller input/output dimensions. Finally, the network has an Average Pooling layer followed by a fully connected layer having 1000 neurons (Image output).

V. Dataset Exploration

RADIOGRAPHY DATABASE

A team of researchers from Qatar University, Doha, Qatar and the University of Dhaka, Bangladesh along with their collaborators from Asian Countries in collaboration with medical doctors have created a database of chest X-ray images for COVID-19 positive cases along with Normal and Viral Pneumonia images. In our current release, there are 219 COVID-19 positive images, 1341 normal images and 1345 viral pneumonia images. We will continue to update this database as soon as we have new x-ray images for COVID-19 pneumonia patients.

This developed database of COVID-19 x-ray images from the Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 DATABASE, Novel Corona Virus 2019 Dataset developed by Joseph Paul Cohen and Paul Morrison and Lan Dao in GitHub and images extracted from 43 different publications. References of each image are provided in the metadata. Normal and Viral pneumonia images were adopted from the Chest X-Ray Images (pneumonia) database.

3. RESULT EVALUATION AND ANALYSIS

I. Dataset

We have total 300 sample images, Which Consist In order 100 of COVID-19 and 100 of PNEUMONIA and 100 of NORMAL Classes Chest X-ray images. For better understanding to train and test Deep neural network with this Large dataset we have portioned this dataset into several parts one for Training and second for Testing Purposes**.** Therefore, percentage-wise portioned training and testing dataset Into 3 sets D_1,D_2,D_3 gives chance to evaluate the accuracy and other parameters of Resnet-50

Neural Network.

All samples = 3×2 table

	Label	Count
1	COVID	100
2	NORMAL	100
3	PNEUMONIA	100

Dataset Name	Training set(%)	Testing set(%)
D_1	50	50
D_2	60	40
D_3	80	20

Num_images = 300

Table 1. Dataset Exploration

II. Confusion Matrix

A Confusion Matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known. To Measure the performance of our algorithm on the testing dataset. This matrix also lets us know the correctly classified images and miss-classified images. By observing the confusion matrix, we can Observe the performance of our deep learning Neural Network. There are Some Basic Terms, which are Useful to Interpret Confusion Matrix.

- True positives (TP): These are cases in which we predicted yes (they have the disease), and they do have the disease.
- True negatives (TN): We predicted no, and they don't have the disease.
- False positives (FP): We predicted yes, but they don't actually have the disease. (Also known as a "Type I error.")
- False negatives (FN): We predicted no, but they actually do have the disease. (Also known as a "Type II error.")

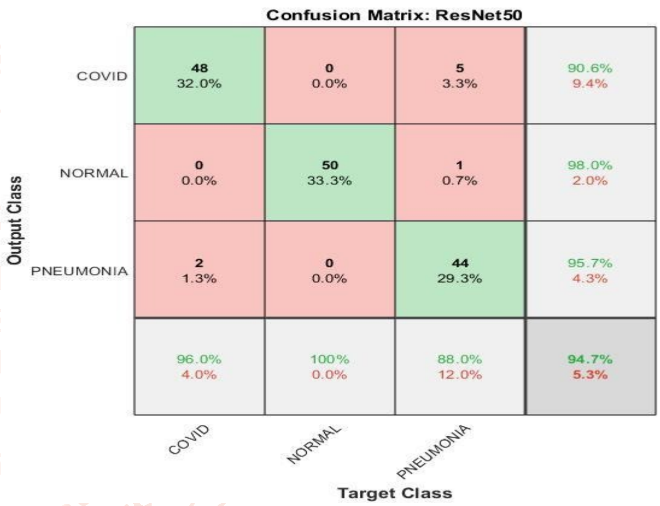


Figure 4. Confusion Matrix

III. Composition and Evaluation

To evaluate and validate the effectiveness of the proposed approach, we

conducted the experiments So many times respectively. Parameters were

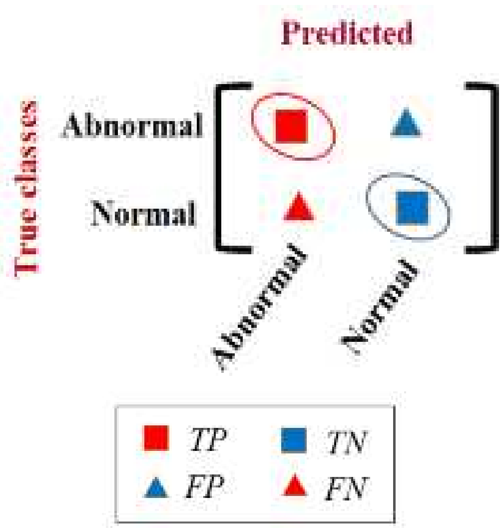
heavily turned to increase the performance of the model.

- For performance evaluation, we adopted Accuracy (ACC), and Sensitivity (SN)metrics from the confusion matrix.

$$Accuracy(ACC) = \frac{TP + TN}{n},$$

$$Sensitivity(SN) = \frac{TP}{TP + FN},$$

- Where, TP is the true positive in case of abnormal and TN is the true negative in case of normal, while FP and FN are the incorrect model predictions for abnormal and normal cases,

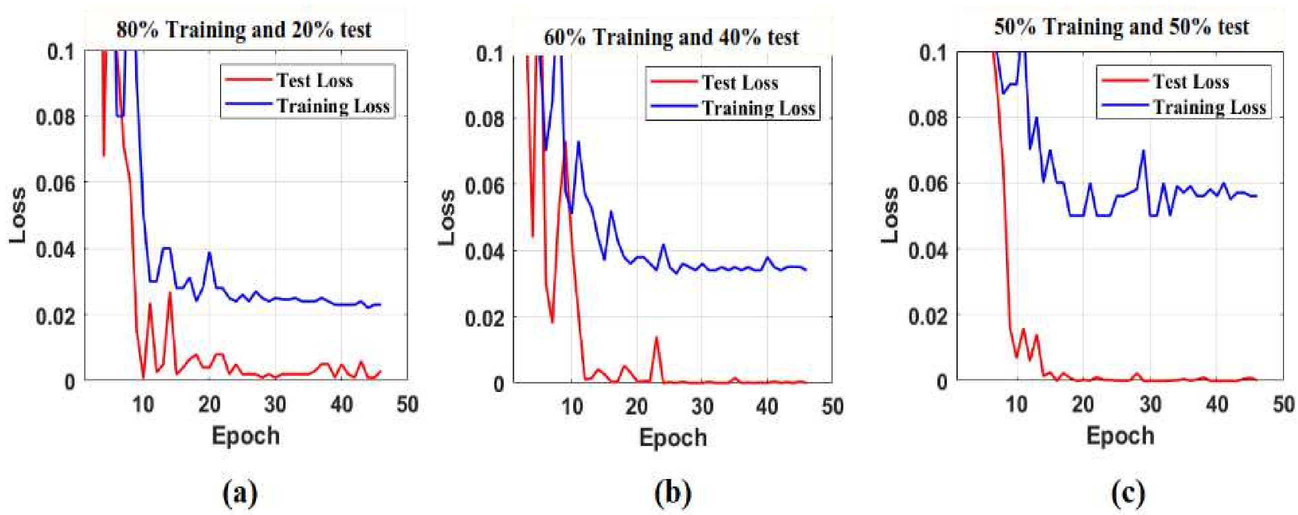


Model	TP	TN	FP	FN
Resnet -50	48	95	5	2

➢ Confusion matrix Tabular For True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) related to the Covid-19 class, for the Resnet-50.

Accuracy = 94.6667	Sensitivity = 0.9600
Specificity = 0.9500	**Precision = 0.9057**

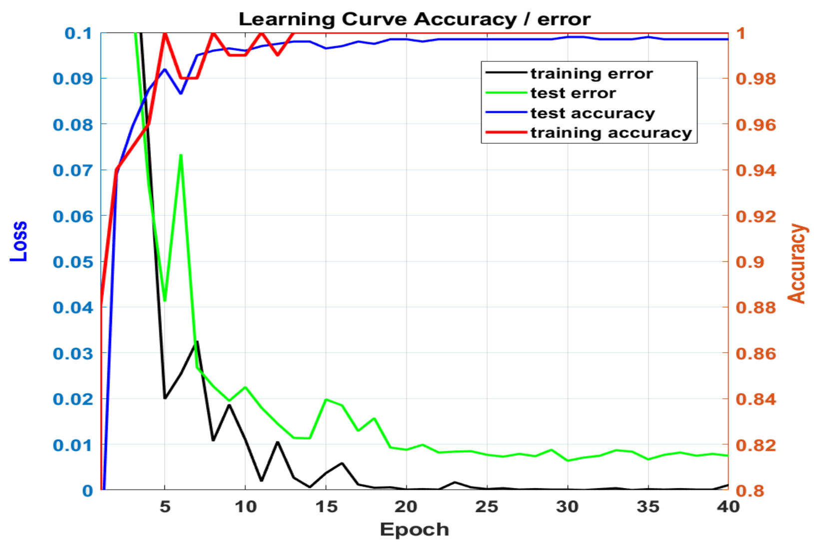
During improvement, a Testing set was used to estimate the model. Fig. is a graphical Representation of the loss value against epoch. In the figures, 'Train loss' indicates training loss, while 'Test loss" indicates validation (testing) loss.



➢ Loss error curves obtained by various sizes for the training set (blue curves) and testing set (red curves): (a) 80% Training and 20% test, (b) 60%Training and 40% test, and (c) 50% Training and 50% test.

IV. Accuracy Graph

Accuracy is a commonly used classification metric and indicates how well a classification algorithm can discriminate the classes in the test set. The accuracy can be defined as the proportion of the predicted correct labels to the total number (predicted and actual) of labels. We have Achieved Highest Accuracy During Training and Testing phase of the Proposed Resnet-50 Neural Network.



The learning curve accuracy and error obtained by Resnet50- Neural network.

- **Activation of a Linear Rectified Unit:**

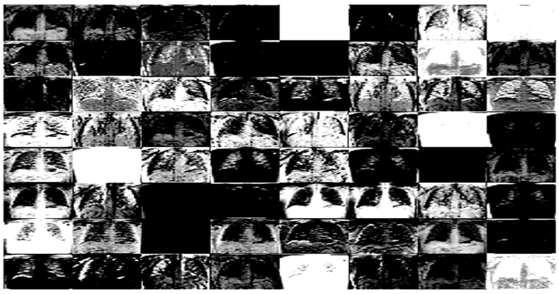
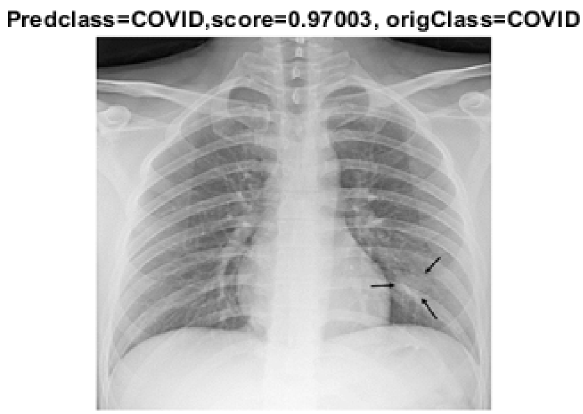
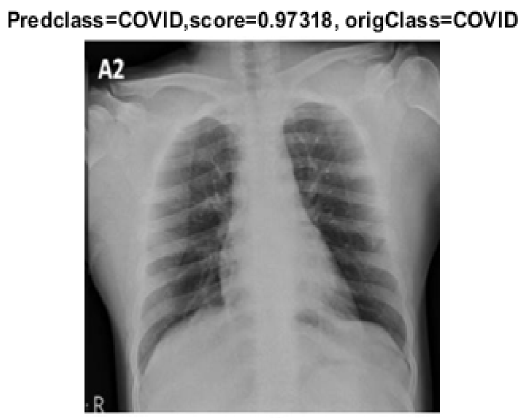


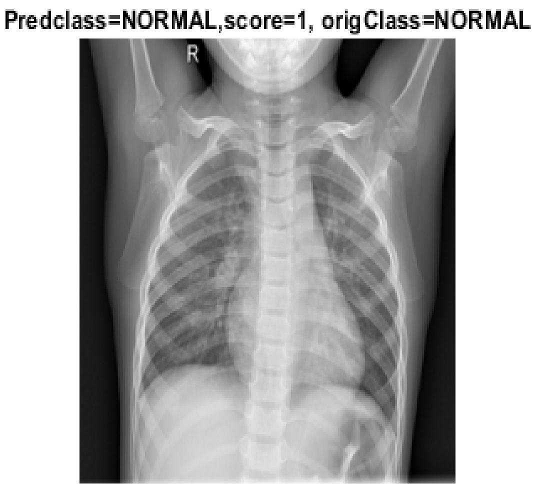
Image Activations From 'activation_2_relu' Layer

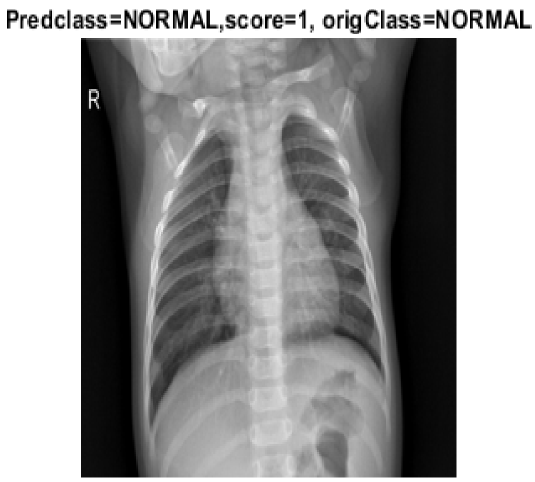
V. Classification Output

- Series of experiments with different proportion of sample datasets have been conducted to measure the effectiveness and accurateness of the proposed system. To ensure the correctness of the knowledge base and outputs, Neural Network is trained repeatedly and its parameters are modified where is needed. For Testing Phase: Loaded Several Images from the test dataset and classify into Three Different Class Using Trained Resnet-50 Neural Network. Results show that accurate outputs, with 94% training accuracy, are resulted with the trained samples, while, the system produces acceptable results for non-trained samples, with 98% classification accuracy.



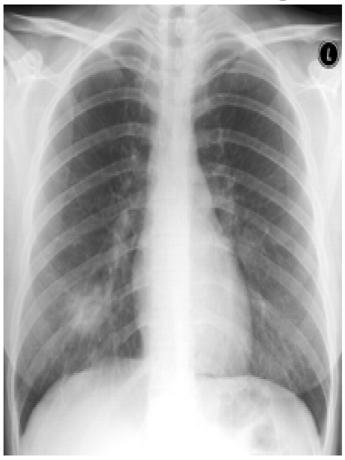






Predclass=PNEUMONIA,score=0.98156, origClass=PNEUMONIA

Predclass=PNEUMONIA,score=0.96443, origClass=PNEUMONIA



Validation (Test) Output

4. CONCLUSION

The creation of medical diagnosis systems is a problem and that has been studied since the early ages. Several techniques and technologies have been used in this field, including both knowledge representation tools and algorithms that perform the diagnosis. Most of the approaches are based on the creation of expert systems which capture the knowledge of a set of medical doctors in order to create a clinical decision support system.

An innovative decision support system which is based on Neural Network and Machine Learning & Deep Learning based Approach has been implemented and tested to classify two main respiratory disease: Covid-19, Pneumonia. Furthermore, the system offers a promising opportunity to develop an operational screening and testing device for heart disease diagnosis. Also, the system can deliver great assistance for clinicians and radiologists to make advanced diagnosis respiratory disease.

Using Patient Xray dataset, series of experiments have been conducted, to examine the performance and accuracy of the proposed system. Compared results revealed that the system performance and accuracy are excellent, with a heart diseases classification accuracy of 98%, results show that the proposed system delivers higher classification accuracy for both respiratory diseases.

More work is planned to enrich the Neural Network and Machine Learning & Deep Learning based Approach knowledge base with more disease samples and enlarge the system to cover a wide range of respiratory diseases with high accuracy. Additional Neural Network training techniques can be adopted, including linear vector quantization (LVQ), perception and dynamic Neural Network. The approach has been validated through the case study; it is possible to expand the scope of modeled medical knowledge. Furthermore, in order to improve decision support, interactions should be considered between the different medications that the patient is on.

REFERENCES

[1] S. Kumar, "Diagnosis of Heart Diseases Using Fuzzy Resolution Mechanism," Journal of Artificial Intelligence, Vol. 5, No. 1, 2012, pp. 47-55. doi:10.3923/jai.2012.47.55

[2] R. A. Miller, H. E. Pople Jr. , and J. D. Myers, "Internist Ian experimental computer based diagnostic consultant for general internal medicine," The New England Journal of Medicine, vol. 307, no. 8, pp. 468–476, 1982.

[3] Q. K. Al-Shayea, "Artificial Neural Networks in Medical Diagnosis," International Journal of Computer Science Issues, Vol. 8, No. 2, 2011, pp. 150-154.

[4] D. L. Hunt, R. B. Haynes, S. E. Hanna, and K. Smith," Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review," Journal of the American Medical Association, vol. 280 no. 15, pp. 1339–1346, 1998.

[5] E. S. Berner, R. S. Maisiak, C. G. Cobbs, and O. D. Taunton, "Effects of a decision support system on physicians' diagnostic performance," Journal of the American Medical Informatics Association, vol. 6, no. 5, pp. 420–427, 1999.

[6] M. L. Graber and A. Mathew, "Performance of a web-based clinical diagnosis support system for internists," Journal of General Internal Medicine, vol. 23, Supplement 1, pp. 37–40, 2008.

[7] E. S. Berner, G. D. Webster, A. A. Shugerman et al. , "Performance of four computer-based diagnostic systems," The New England Journal

of Medicine, vol. 330, no. 25, pp. 1792–1796, 1994.

[8] H. Yana, Y. T. Jiangb, J. Zhenge, C. Pengc and Q. Lid. “A Multilayer Perceptron-Based Medical Decision Sup- port System for Heart Disease Diagnosis,” Elsevier Ex- pert Systems with Applications, Vol. 30, No. 2, 2006, pp. 272-281. doi:10. 1016/j. eswa. 2005. 07. 022

[9] R. Das, I. Turkoglu and A. Sengur, “Diagnosis of Heart Disease through Neural Networks Ensembles,” Computer Methods and Programs in Biomedicine, Vol. 93, No. 2, 2009, pp. 185-191. doi:10. 1016/j. cmpb. 2008. 09. 005

[10] Rahib H. Abiyev , Mohammad Khaleel Sallam,” Deep Convolutional Neural Networks for Chest Diseases Detection", Journal of Healthcare Engineering , vol. 2018, Article ID 4168538, 11 pages, 2018. https://doi.org/10.1155/2018/4168538

[11] Aronson, “Diagnosis Pro: the ultimate differential diagnosis assistant,” Journal of the American Medical Association, vol 277, no. 5, p. 426, 1997.

[12] DeTraC: Transfer Learning of Class Decomposed Medical Images in Convolutional Neural Networks, April 2020, IEEE Access PP(99); **Authors:** Mohammed Abdelsamea, Mohamed Medhat Gaber DOI: 10.1109/ACCESS.2020.2989273,

[13] Zak M. , Krzyżak A. (2020) Classification of Lung Diseases Using Deep Learning Models. In: Krzhizhanovskaya V. et al. (eds) Computational Science – ICCS 2020. ICCS 2020. Lecture Notes in Computer Science, vol 12139. Springer, Cham. https://doi.org/10.1007/978-3-030-50420-5_47

[14] Classification of Lung Diseases Using Deep Learning Models, Authors: Matthew Zak, Adam Krzyżak, Published On 15 June 2020.

[15] CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, **Authors:** Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Robyn L. Ball, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. Ng. Published on 25 Dec 2017.

[16] A Computer Aided Diagnosis of Lung Disease using Machine Learning Approach, **Authors:** Subapriya V, Jaichandran R, Shanmugaratnam K. L, Abhiram Rajan, Akshay T, Shibil Rahman. Published in April 2020.

[17] Convolutional Neural Network Based Classification of Patients with Pneumonia using X-ray Lung Images, Published in May 2020, **Authors:** Hicham Moujahid, Bouchaib Cherradi, Oussama El Gannour, Lhoussain Bahatti, Oumaima Terrada, Soufiane Hamida, DOI:10.25046/aj050522.

[18] S. Mehrabi, M. Maghsoudloo, H. Arabalibeik, R. Noor- mand and Y. Nozari, “Application of Multilayer Percep- tron and Radial Basis Function Neural Networks in Dif- ferentiating between Chronic Obstructive Pulmonary and Congestive Heart Failure Diseases,” Expert Systems with Applications, Vol. 36, No. 3, 2009, pp. 6956-6959. doi:10.1016/j.eswa.2008.08.039

[19] P. Baldi, S. Brunak, Y. Chauvin, C. A. F. Andersen, and H. Nielsen, “Assessing the accuracy of prediction algorithms for classification: an overview,” Bioinformatics, vol. 16, no. 5, pp. 412–424, 2000.

[20] M. J. Lincoln, C. W. Turner, P. J. Haug et al. , “Iliad training enhances medical students’ diagnostic skills,” Journal of Medical Systems, vol. 15, no. 1, pp. 93–110, 1991.

[21] Rodr´ıguez-Gonz´alez, J. E. Labra-Gayo, R. Colomo-Palacios, M. A. Mayer, J. M. G ´omez-Berb´ıs, and A. Garc´ıa-Crespo, “SeDeLo: using semantics and description logics to support aided clinical diagnosis,” Journal of Medical Systems, vol. 36, no. 4, pp. 2471–2481, 2012.

[22] G. O. Barnett, J. J. Cimino, J. A. Hupp, andE. P. Hoffer, “DXplain: experience with knowledge acquisition and program evaluation,” Proceedings of the Annual Symposium on Computer Application in Medical Care, no. 4, pp. 150–154, 1987.