
# LUNG NODULE DETECTION IN CHEST X-RAY USING DEEP LEARNING

***Kamran Javaid*[1], *Dr. Faisal Rehman [1], Dr. Nuhman-ul-Haq[1]***
*[1]Computer Science Department COMSATS University, Islamabad Abbottabad Campus Pakistan*
*kamranbajna@gmail.com*

## ABSTRACT

Lung disease is a type of pulmonary disease. It is estimated that currently there are 600 million pulmonary patients worldwide and by the year 2028 the said malady will likely become the third largest fatal disease. Modern diagnostic technologies are available for diagnosis of pulmonary diseases, but due to limited access to modern technologies and the financial burden on the rural population, X-ray is an effective technique to deal with pulmonary diseases. Lung cancer is also a pulmonary disease, which is the world's most severe cancer. Most of the time patient is unaware of lung cancer until an X-ray is performed. The manual analysis of the X-ray image is not only expensive, but error-prone. Therefore, various Computer-Aided Diagnosis (CAD) based techniques have been applied to automatically detect the cancer nodule present in the X-ray image. Recently, Deep Learning based Convolution Neural Network techniques have been used to detect the diseases present in X-ray images. However, deep learning systems need a huge number of images for training before they are implemented in the real-world scenarios and the deployment of these systems are still in the initial stages as it is extremely challenging to automatically recognize or detect cancer present in an X-ray image. The problem arises when we face a lack of training images due to which the performance of such systems decreases exponentially. This research focuses to use the balanced dataset with Transfer-Learning based Deep Learning network by combining Vgg16 and Custom CNN to classify model with high accuracy, our model yields 88% of Accuracy, 83% of Sensitivity, and 92.4% of Specificity in Binary class while in Multi-class we have achieved 89% of Accuracy, 89% of Sensitivity and 94% of Specificity. The proposed study is performed both in Binary class and multi-class & our proposed research has outperformed the current cutting-edge models.

**Keywords—** Transfer-Learning; Lung Nodule; X-ray; Convolution Neural Network; Deep Learning.

## 1. INTRODUCTION

The lung is the most important part of the human body which takes oxygen to the body and removes carbon dioxide from the body. With rapid development in the Industrial sector, air pollution, and transportation, the incidents of lung diseases have increased rapidly. Worldwide, there are 600 million patients with pulmonary disease, and by 2028 it is expected to become the third deadly disease [1].

Lung cancer is also one of pulmonary disease which occurs due to the uncontrolled proliferation of tissues in the lungs [2]. A pulmonary nodule is a lung disease that shows certain symptoms like coughing up blood, chest pain, weight loss, and shortness of breath [3].

Health Informatics is the scientific study that converges digital technologies and healthcare. Medical informatics incorporates two types of medical data when conducting a field study; those are the data from biomedical analysis and imaging data, which have unique characteristics. Following the imaging modalities, Pixels that refer to a part of a physical object to produce digital image data & biomedical records are created from diagnostic examinations of patients. The characteristic difference between biomedical recording and the image data is what makes the methods and procedures required to explore them differently. Medical image data are produced using modalities for imaging. In this area, the question is how to extract the image and classify the extracted product into a similar pattern, then identify and understand the parts of the human body that are affected by the disease from the image classification product. Medical image classification is presented in three phases: (1) Image feature extraction and description, (2) Collection of features to be used for classification and (3) classification of images based on features.

An image is deemed to represent an object with key features used in image processing. Medical image processing activities include the extraction and interpretation of objects. Biomedical Equipment produces the medical image data, using imaging techniques such as CT, Mammograms, and Magnetic resonance imaging scans. There is a range of medical imaging modalities which include ionizing radiation, nuclear medicine, magnetic resonance, ultrasound, and optical methods as modality media. Each of the modality media has unique characteristics and differences in the structure of the human body and the organ tissue.

Computer-aided Diagnose (CAD) based techniques have been applied to automatically detect the diseases present in the image. Pulmonary nodule identification, segmentation, and classification are the main functions of the lung nodule-based CAD system and are part of computer vision. The hands-on engineering methods are unable to deal with complicated features and hence lead to poor classification. The X-ray procedure is one of the vast majorities of the least expensive open radiology assessments. They are commonly the best option of specialists and radiologists as a result of its minimal effort and a low portion of radiation and its capacity to uncover data that may regularly go unnoticed that is why X-ray is commonly used as a diagnostic method. With the number of patients increasing, the radiologists are facing workload. This makes it essential to develop an automated and computational way of understanding the content of CXR images. In the area of computer vision, deep learning has recently revolutionized. Deep learning has shown promising results in various fields in the past few years, several deep learning-based models are used for Medical image diagnosis, however, these models are hard to train because they require a tremendous amount of data to train to perform accurately, the main issue related to deep learning-based disease diagnosis systems arises when we are having a very little amount of training sample. In particular deep learning classifiers, CNN is widely used for image classification problems in recent years. Currently, CNN-based architectures are primarily used to identify and classify diseases present in an image. A Deep Learning based CAD system for image classification is presented in Fig. 1.

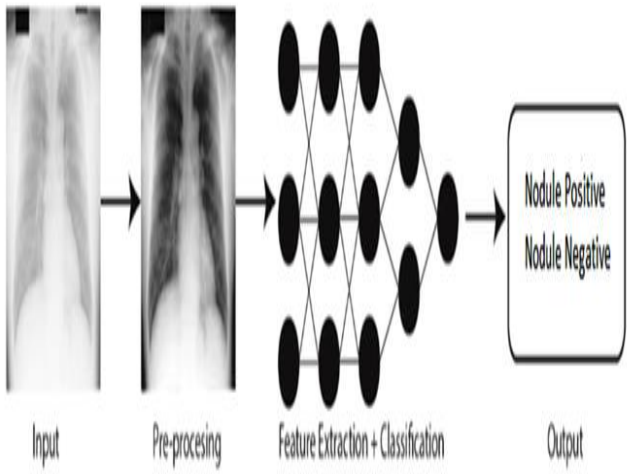


*Figure 1: Deep Learning-Based CAD System*

The estimated patients with chronic obstructive pulmonary diseases are 600 million worldwide and it is reported that by 2028 it will be the third fatal disease. During the Chest X-ray examination, about 20-50 percent of lung nodules are either misdiagnosed or missed by the radiologist [1]. As reported by WHO, 8.8 million deaths were caused by cancer in 2015, out of which 1.69 million were due to lung cancer and 70% of these deaths occur in low and middle-income countries [4].8.8 million Deaths were caused by cancer in 1 year, out of which 1.69 million were due to lung cancer. Chest X-ray is the most widely accepted & easily available technology and generally, it's the first choice of medical specialists because of the ability to disclose information that can sometimes go unreported. X-ray technology is important because of its low radiation, pathological alternation, and low cost. Lung cancer is the most common cancer in men and the second most common cancer in women in the world. Figure.2 & Figure.3 shows the lung cancer cases & death caused by lung cancer in 2018. According to a study by Globocon, 2012, In Pakistan lung cancer is the third most common cancer [5]. With a growing increase in the population of patients, radiologists facing an increasing workload, this makes it important to build an automated way of understanding CXR images with effective performance.

Nodule screening in cancer treatment plays a significant role in preventing consideration as it is mostly curable if caught in the early stages. Recently reported by the Care Quality Commission that in the UK throughout a 1-year, the total number of 23,000 CXR images was not properly reviewed at Queen Alexandra Hospital alone [6]. Queen Alexandra Hospital doesn’t just have the issue of providing an expert reading for every chest X-ray. The growing population increase demand for CXR reading. Planning and delivery of diagnosis of cancer diseases are not only complex, but cost-effective which can be facilitated by Deep learning & Machine Learning, which are the fastest-growing fields in Artificial Intelligence which are successfully deployed in recent years in many domains including medical, health and agriculture, etc.

The research of CAD system has been concerned by many data scientists and it has been applied for disease classification e.g. pulmonary disease classification, however, there are still many defects and problems that should be considered. In this research, we aim to improve the efficacy of the CAD system in the diagnostic role of the lung nodule. Deep-Learning based Convolution Neural Network techniques have been used to detect the diseases present in X-ray images. However, deep learning systems need a huge number of images for training before they are implemented in the real-world scenarios and the development of these systems is still in the early phase as it is incredibly difficult to automatically classify cancer present in Pulmonary X-ray. The problem arises when we face a lack of training images due to which the performance of such systems decreases exponentially. It is found from the literature that no one has addressed multi-class nodule classification in the JSRT dataset for classifying nodule type, so we are addressing both Binary and multi-class classification of lung nodules in the JSRT dataset with high accuracy.

### 1.1 RELATED WORK

In the year 2019, authors have compared Artificial Neural Network (ANN) with traditional machine learning techniques for X-ray classification and found that ANN outperformed traditional techniques with an accuracy of 83%, sensitivity of 82.35%, and specificity of 82.61% on the JSRT dataset after pre-processing is applied but authors only discussed the presence and absence of nodule but no information about the type of nodule is provided and no noise removal is performed on the dataset [2]. In 2019, the author proposed the study of the classification of respiratory images based on the Inception-v3 transfer learning model, and the proposed work was compared to the previous work, DCNN is used as a feature extractor, and classification was done using SVM, Logistic, and SoftMax classifiers. The author found that the neural network model based on Transfer-learning neural network model performs better in JSRT than the original DCNN model. The performance during binary classification, resulting in a Sensitivity of 95 %, Specificity of 80 % & Accuracy of 86 % on the JSRT dataset, but the author only discusses the presence and absence of a nodule, and no information is provided about the nodule type [7]. In 2019, the author proposed the Faster regional CNN model and classified chest X-ray image into 2 categories, Pathological and Normal based on a subset of the Xray-14 dataset. To boost its precision in object detection, the area proposal network was integrated into CNN. It was then improved in terms of time execution by combining multiple layers into one network followed by the elimination of the area proposal extraction earlier, but the author has not provided any information about the pathological disease and nodule type. The size of the dataset was very small and the author has used a random subset of the dataset which will lead toward biasness [8]. In 2019, Three CNN's are studied and the impact of layers on the classification of X-ray images was studied. The scale of the initial X-ray chest images was of size 1024×1024. If this size of images were fed into the networks, the classification performances will be greater. However, using such a large-size input image will take a very long-time during learning. It is found that classification efficiency for X-ray chest images does not decrease significantly with the input image of size 128×128, classification performances for the X-ray chest images do not decrease dramatically. The chest image is assigned to 12 classes, including Nodule, Edema, etc. With a success rate of 86% on the Chest X-ray14 dataset, but nodule type is not specified, only the presence and absence of nodule is discussed in the study [9].

In 2018, the author proposed a partial solution by integrating multi-resolution and multi-instance learning with a personalized pooling feature designed to provide more precise diagnosis and pathology-based, higher-resolution gray level charts. The findings are not very much promising which indicates the need for a potential study where the proposed model may be pre-trained to enhance output on data from another area. [10].

## 2. MATERIAL AND METHODS

### 2.1 DATA AND PREPROCESSING

Data augmentation is essential in lung image nodule classification as there are shortcomings of medical images in lung nodule detection. Medical image labeling cannot be outsourced, unlike voice data or textual data. The relevance of data augmentation lies in the fact that the amount of training data can be reasonably and fairly increased on the limited pulmonary image data set to improve the model's generalization ability. The technique augmentation is used to expand the datasets to avoid overfitting as datasets have fewer images furthermore contrast of all images is increased to normalize the intensity of the images. After augmentation, the number of images for each class is expanded up to 2500.

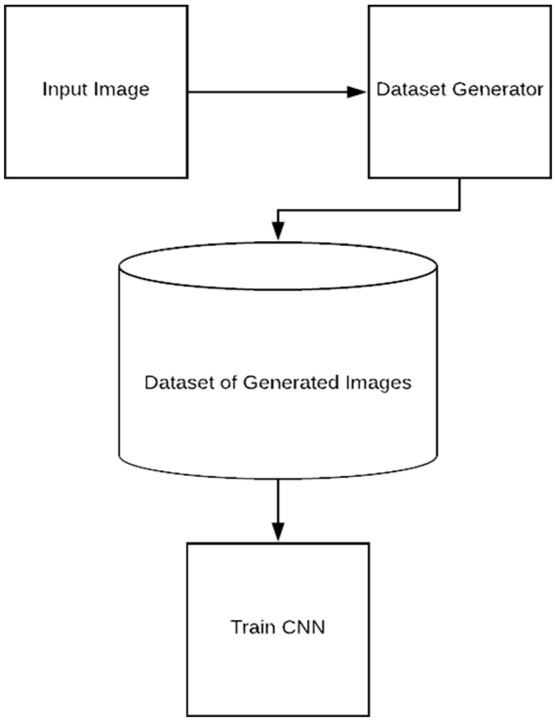


*Figure 2 Data Augmentation*

## 2.2. PROPOSED MODEL

We build two models, one for binary classification and the other for multi-class classification, both the model architectures were the same and based on transfer learning. VGG-16 comprises of **sixteen convolutional layers** and is very attractive & acceptable because of its **uniformity.** The model's input image is of measurements (224, 224, 3). Vgg-16's initial two layers have 64 channels, with 3 * 3 filter size and the same padding. Then after a max pool stride layer of size (2,2), two layers that have 256 filter size and filter size convolution layers of (3,3) are there. This was followed by a max-pooling stride layer (2, 2) that is the same as the previous layer. Then there are 2 filters-size (3, 3) and 256 filter convolution layers. Following that there are two sets of 3 layers of convolution and one layer of the max stream. Each has 512 filters of (3,3) size with the same padding. Then this image is passed to the heap with two convolution layers. The filters we use in those layers of convolution and max-pooling are of the size 3*3. It also uses 1*1 pixels in some of the layers which are used to control the number of input channels. Figure 7 shows the architecture of Vgg16 network.

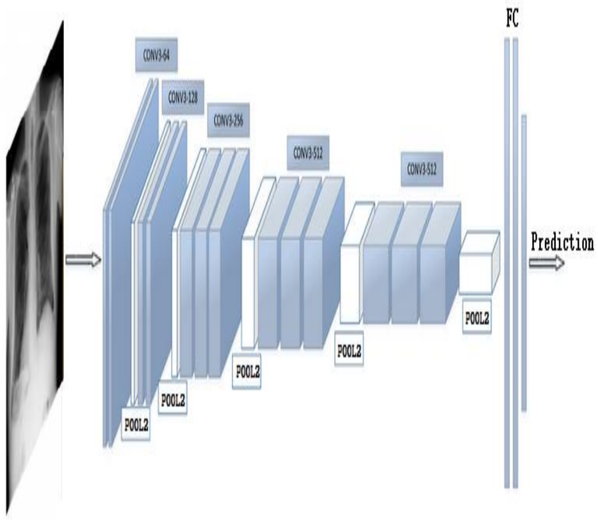


*Figure 3: Vgg16 Architecture*

Increasing the size of the convolutional neural network would also increase network performance, but this move comes at the price of utilizing resources. Thus, to reduce the training cost, a Transfer Learning network arises. Transfer learning is to transfer the parameters of the learned network to the new model in order to assist the new training model. Assuming that certain data or tasks are important, the learning efficiency performance of the new model can be improved and optimized by sharing the training model's parameters with the new model. By passing on information rather than starting from scratch as most networks do.

We combined two models as shown is fig. 4, One we pick the famous transfer learning model called Vgg-16, and second we build custom CNN, we set the two layers of VGG as trainable, those two layers are: **'block5_conv1', 'block4_conv1'**, and extract bottleneck features from it and the output of VGG becomes an input for CNN, CNN was consisted of 4 layers, the first one is the input layer and the last one is the output layer, and between these input-output layers, there were 2 hidden layers consist of 512 neurons.

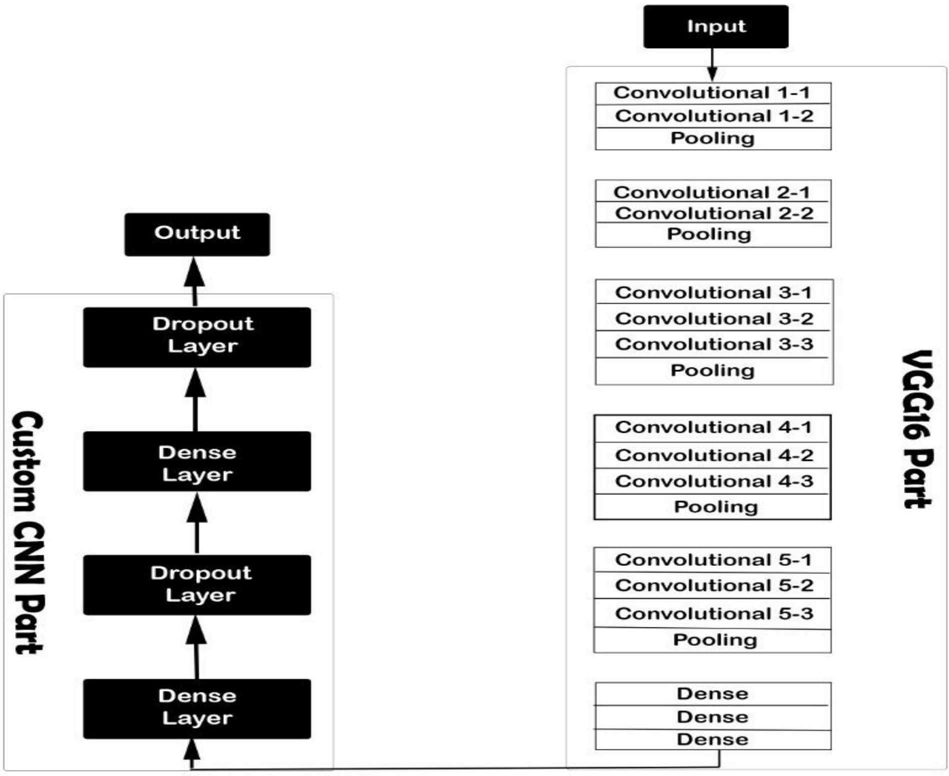


*Figure 4: Model Architecture*

## 3. EXPERIMENTAL RESULTS AND ANALYSIS

The efficiency of the proposed study is assessed through its effect on chest x-ray detection of nodules.

We pre-processed the data, prepared it for the network's input, and trained the model numerous times so that it could eventually learn to identify the disease. This can be used as a method of doctor verification in a hospital management setting.

A comparison with other existing methods has been provided in the study. A Transfer-Learning based Deep-learning network was formulated on the dataset to predict nodule presence or absence along with the type of nodule. Various training and validation runs were performed on Chest X-rays images from the JSRT dataset. Model performance in the research was assessed using accuracy, sensitivity, and specificity. Accuracy indicates to what degree the model did identify all the positive and negative cases correctly. Specificity refers to what degree the model did identify negative cases correctly. The sensitivity indicates to what degree positive cases are observed correctly in the model. By large value of accuracy, specificity, and sensitivity, it can be inferred that the network has low error potential.

### *3.1 BINARY MODEL*

The performance of the proposed model is assessed by its effect on chest x-ray detection of nodules. Also provides a comparison with other existing methods. The model had been trained on the JSRT dataset to predict nodule presence or absence.

There are 2 classes in the Binary model. Class 0 & Class 1. Class 0 indicates the Nodule category and Class 1 indicates Non-Nodule. Binary classification comprises of 5087 images, out of which we used 4887 images for model training and validation and 200 used in model testing.

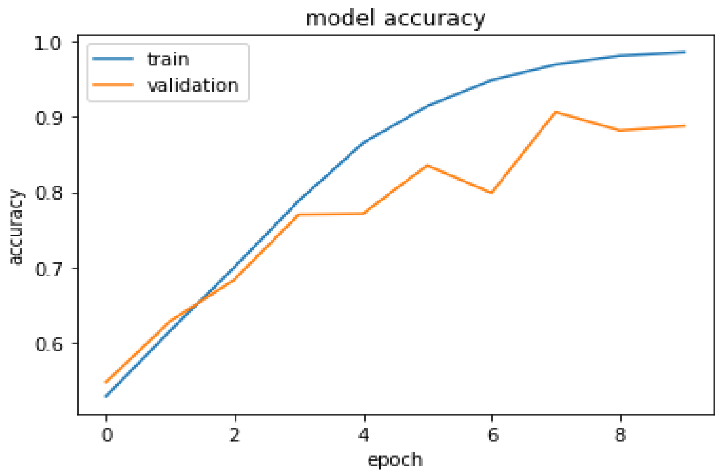


*Figure 5: Model Accuracy*

Fig. 5 shows the Training and validation accuracy of the Binary model in which the diagram showing good results on both training and validation datasets. The training dataset is for model learning, while the validation dataset validates the generalization ability of the model for early termination during training the model.

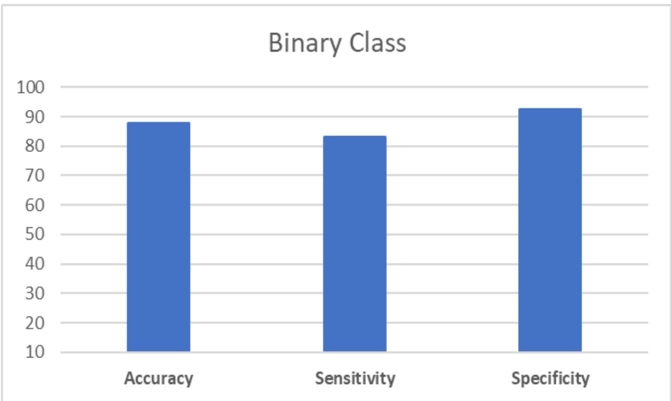


*Figure 6: Overall Performance Binary Model*

Fig. 6 shows the Accuracy, Sensitivity & Specificity of the Binary-class model. Our model yields 88% of Accuracy, 83% of Sensitivity, and 92.4% of Specificity in Binary class. Accuracy shows Positive results truly predicted by our model; Specificity represents the probability of correctly classifying nodule negative & Sensitivity value corresponds to the probability of correctly classify nodule positive images.

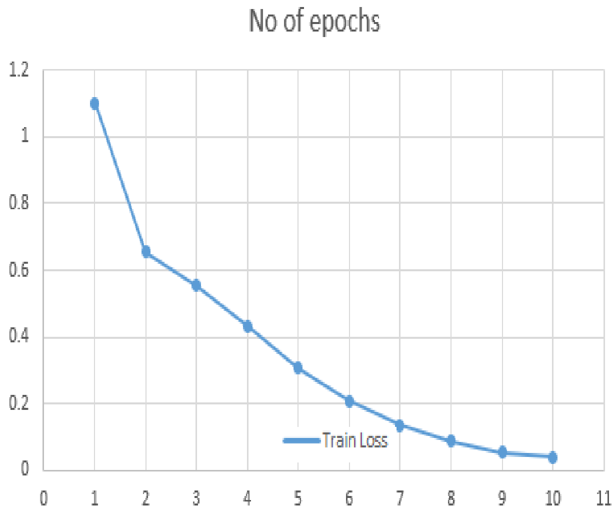


*Figure 7: Training Loss Binary Class*

Fig.7 shows the Training loss of the Binary-class model, which on 1st Epoch was 1.1, and with epoch by epoch by training of our model on the dataset we got the training loss of 0.038 after 10 epochs. The value of training loss should be as small as it is possible.

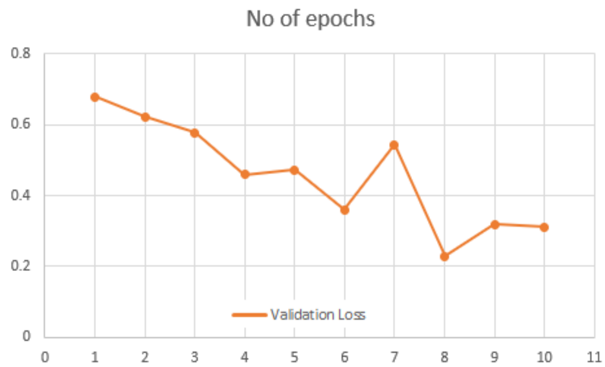


*Figure 8: Valuation Loss Binary Class*

Fig. shows the Validation loss of the Binary model. If in an experiment validation loss going to increase that means model is going toward overfitting. We set the numbers of epochs as high as possible and to avoid overfitting. After 10 epochs the value of validation loss decreased to 0.310 which is a reasonable loss with desired performance.

### *3.2.MULTI-CLASS MODEL*

The efficiency of the proposed model is measured by its effect on the identification of the nodule type in chest x-ray and by comparing it with other cutting-edge models. The network model had been trained for predicting the presence (Malignant or Benign Nodule) or absence (Non-Nodule) of the nodule on the JSRT dataset.

There are 3 Classes in Multi-class classification. Class 0, 1, and 2. Class 0 represents the benign cases, class 1 represents Malignant and Class 2 represents the Non-nodule image types. The total number of images in Multi-class classification was 7365, out of which we used 6965 images for training & validation and 400 for testing purposes.

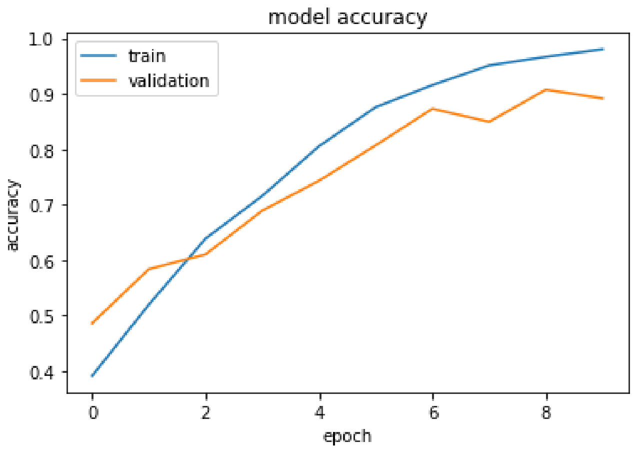


*Figure 9: Model Accuracy- Multi-class*

Fig. 9 shows the Training and validation accuracy of the Multiclass model. Showing good results on both training and validation data sets. Training accuracy is related to the training dataset, the one which is used to fit the model, while validation accuracy is related to the validation dataset, which is used to validate the generalization ability of our model for early termination during model training.


*Figure 10: Overall Performance- Multiclass Model*

Fig. 10 shows the Accuracy, Sensitivity & Specificity of the Multi-class model. These are the most effective scientific criteria for assessing the image in the Medical field. The greater the value of sensitivity is, the less the frequency of missed diagnosis in the model; the greater the value of specificity is, the less the risk of wrong diagnosis rate is there. However, the specificity can decrease as sensitivity increases, so it is important Sensitivity and specificity weigh high. The findings of this experiment's sensitivity, specificity, and accuracy are relatively high.

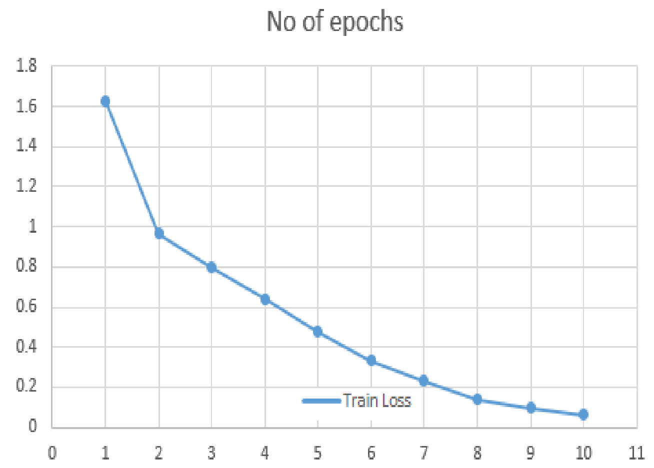


*Figure 11: Training Loss Multi-class model*

Fig. 11 shows the Training loss of the multi-class model, which on 1st Epoch was 1.6, and with epoch by epoch, with the training of our model, we got the training loss of 0.062 after 10 epochs. The value of training loss should be as small as possible.

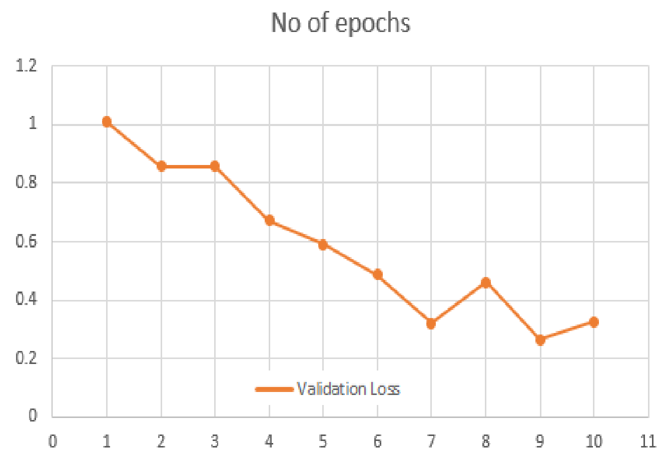


*Figure 12: Validation Loss- Multiclass model*

Fig. 12 shows the Training loss of the multi-class model. If in an experiment validation loss going to increase that means model is going toward overfitting. We set the numbers of epochs as high as possible and to avoid overfitting. After 10 epochs the value of validation loss decreased to 0.325.

**Table 1. Different models and their accuracy**

| **Model** | **Accuracy** |
|---|---|
| [2] Artificial Neural Network (ANN) | 82% Binary Class |
| [7] Inception v3 & SVM | 86% Binary Class |

| | |
|---|---|
| [11] Convolutional Neural Network (CNN) | 82% Binary Class |
| [12] Faster Regional CNN Model | 62% Binary Class |
| [13] Res Net 18 | Avg Accuracy = 84%, Nodule class, 79% |
| **Our Model (Vgg16+ Custom CNN)** | **88%** Binary Class, **89%** Multi Class |

### *3.3. NETWORKS COMPARISON*

The network models have been trained to predict nodule presence or absence along with the type of nodule type (Malignant or Benign). Several Training and validation runs were performed on the dataset containing X-ray images. The confusion matrix is used which is the most widely used tool for assessing the effectiveness of the systems obtained from data sets with pre-determined target data in the field of deep learning and machine learning. We evaluated the performance of our classification models using a number of assessment criteria's which are often used in previous research to compare our models with previous work as shown in Table III, we created a confusion matrix for each model and evaluated Accuracy, Sensitivity & Specificity from it.

There is a significant improvement for nodule classification in terms of Accuracy, Specificity & Sensitivity. We can conclude that the Transfer-Learning based deep learning model can perform better in our classification problem as our network models achieved 88% of Accuracy, 83% of Sensitivity, and 92.4% of Specificity in Binary class while in Multi-class we have achieved 89% of Accuracy, 89% of Sensitivity and 94% of Specificity which is higher than models reported in the literature.

In comparison to deep learning methods, Transfer-Learning based deep learning model has better performance in terms of performance metrics. Though its effect is slightly lower than some hand-crafted extraction techniques. However, CNN can auto extract features to save considerable time and effort. On the JSRT X-ray images dataset, we have presented a review of different models of X-ray image classification done through Convolutional neural networks. Although the performance results are satisfactory with networks configured on the ImageNet dataset. Our results showed that the method of lung nodule classification in X-ray images based on transfer learning performs better.

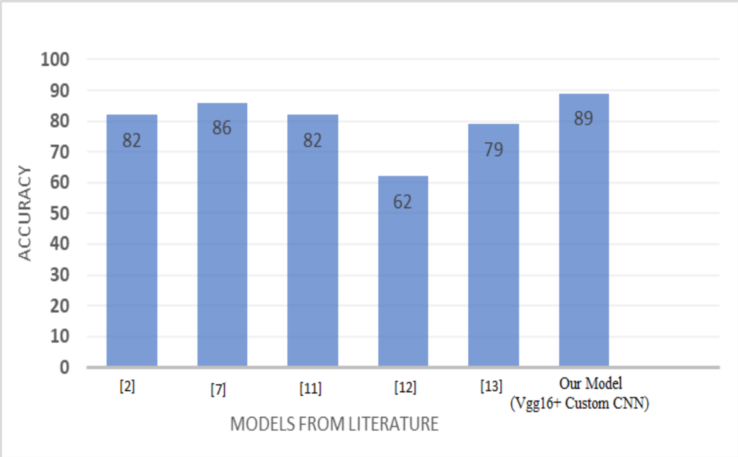


*Figure 13: Models Comparison*

The model's Accuracy value is compared to other methods as shown in Table 1 & Fig. 13 and it is noted that our framework does better than the other methods previously presented in the literature over the last 2 years.

On the basis of Vgg-16 transfer learning in Chest X-ray images, we presented an approach for classifying lung pictures. Higher accuracy can be attained by using the transfer learning-based lung image categorization technique. Additionally, the neural network model based on transfer learning outperforms the model based on models designed from scratch in the classification of lung images on the Chest X-ray image database. Accuracy can be significantly increased by fine-tuning transfer learning.

## 4. Conclusion

The automatic image classification procedure of lung nodules is presented in this paper and compared with previous work. The number of images was very few so for better accuracy, pre-processing and data augmentation techniques are performed. We build two models, one for binary classification and the other for multi-class classification; both the model architectures were the same and based on transfer learning. Two models were combined. First, we choose the well-known transfer learning model Vgg-16, and then we created a custom CNN model. It has been observed that training a model from scratch is not an effective approach while network initialization with transfer learning performs well.

A publicly available JSRT dataset is utilized for training and performance assessment of the networks. The model's performance is assessed using Accuracy, Sensitivity & Specificity. It achieves 88% accuracy, 83% sensitivity, and 92.4% specificity in the binary class and 89% of accuracy, 89 % sensitivity, and 94% specificity in multi-class classification. The model's performance also outperforms most of the methods reported in the literature. Our Transfer learning based deep learning model shown some outstanding results.

There is a wide difference among specificity and sensitivity, Therefore, in future research work, a study can be carried out to minimize the gap between specificity and sensitivity & this solution is only best suited for pulmonary tasks and is not designed for all other medical data so that more studies can be carried out on other medical images for further research.

The proposed system can be used as initial screening but integrating this into a user-friendly Software application will be helpful for non-technical person to apply in the real-world situation.

## 5. REFERENCES


[1] [M. A. Bruno, E. A. Walker, and H. H. Abujudeh, "Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction," Radiographics, vol. 35, no. 6, pp. 1668–1676, 2015.

[2] Ö. Günaydin, M. Günay, and Ö. Şengel, "Comparison of Lung Cancer Detection Algorithms," in 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 2019, pp. 1–4.

[3] J. Shiraishiet al., "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," Am. J. Roentgenol., vol. 174, no. 1, pp. 71–74, 2000.

[4] "Cancer," 31-Dec-2019. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer. [Accessed: 31-Dec-2019].

[5] Sarwar MR, Saqib A. Cancer prevalence, incidence and mortality rates in Pakistan in 2012. [Online] 2017 [Cited 2017 Oct 9]. Available from:URL: https://www.cogentoa.com/article/10.1080/2331205X.2017.1288773

[6] "Queen Alexandra Hospital," 31-Dec-2019. [Online]. Available: https://www.cqc.org.uk/location/RHU03. [Accessed: 31-Dec-2019].

[7] C. Wang et al., "Pulmonary Image Classification Based on Inception-v3 Transfer Learning Model," IEEE Access, vol. 7, pp. 146533–146541, 2019.

[8] R. Hooda, A. Mittal, and S. Sofat, "Automated TB classification using ensemble of deep architectures," Multimed. Tools Appl., vol. 78, no. 22, pp. 31515–31532, 2019.

[9] E. Kesim, Z. Dokur, and T. Olmez, "X-Ray Chest Image Classification by A Small-Sized Convolutional Neural Network," in 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 2019, pp. 1–5.

[10] L. Yao, J. Prosky, E. Poblenz, B. Covington, and K. Lyman, "Weakly supervised medical diagnosis and localization from multiple resolutions," ArXivPrepr. ArXiv180307703, 2018.

[11] M. Uçar and E. Uçar, "Computer-Aided Detection of Lung Nodules in Chest X-Rays using Deep Convolutional Neural Networks," Sak. Univ. J. Comput. Inf. Sci., vol. 2, no. 1, pp. 41–52, 2019.

[12] A. Ismail, T. Rahmat, and S. Aliman, "CHEST X-RAY IMAGE CLASSIFICATION USING FASTER R-CNN," Malays. J. Comput., vol. 4, no. 1, pp. 225–236, 2019.

[13] S. Rakshit, I. Saha, M. Wlasnowolski, U. Maulik, and D. Plewczynski, "Deep Learning for Detection and Localization of Thoracic Diseases Using Chest X-Ray Imagery," in International Conference on Artificial Intelligence and Soft Computing, 2019, pp. 271–282.