# IMPROVED CHEST X-RAY DIAGNOSIS USING DEEP LEARNING FOR MULTI-CLASSIFICATION OF PATHOLOGIES

***Project Reference No.:*** *47S_BE_4590*


***College*** *: Ramaiah Institute of Technology*
***Branch*** *: Department of Computer Science and Engineering*
***Guide(s)*** *: Mr. Pradeep Kumar D*
***Student(S)*** *: Mr. S D Hrudhay*
*Mr. Sajal Srivastava*
*Mr. Samyak Jain*
*Mr. Shashank Singh*


**Introduction:**

Medical imaging, particularly the diagnosis of chest X-rays, is essential for the early identification and management of several thoracic illnesses. On the other hand, prompt and precise diagnosis is becoming increasingly difficult due to the growing amount of medical data. The current issue is the requirement for a scalable and reliable system capable of non-binary i.e., multi-label classification with the aim of identifying different thoracic pathologies from pictures of the chest X-ray. The intricacy stems from the fact that various illnesses are varied and need different approaches for successful identification. Furthermore, the presence of many diseases in one X-ray picture adds another level of complexity that necessitates sophisticated computational solutions. The primary goal is to create an automated system that can accurately classify several labels and identify multiple common thoracic illnesses, such as 'Cardiomegaly', 'Emphysema', 'Effusion', 'Hernia', 'Infiltration', 'Mass', 'Nodule', 'Atelectasis', 'Pneumothorax', 'Pleural Thickening', 'Pneumonia', 'Fibrosis', 'Edema' and 'Consolidation'. The study aims to address the challenges associated with anatomical variability and picture quality by utilizing sophisticated deep learning models and extensive chest X-ray datasets. The significance of this endeavor extends beyond mere technological innovation. A more accurate and efficient diagnostic tool holds the potential to revolutionize patient care by enabling earlier disease detection, facilitating timely intervention, and ultimately improving patient outcomes. Furthermore, by automating and enhancing the diagnostic process, we can alleviate the burden on healthcare professionals, allowing them to focus on more complex clinical decision-making and personalized patient care.

**Objectives:**

1. Develop a Probabilistic Model from comparative analysis of various Deep Learning Models to Enhance Pathology Identification in Chest X-ray Medical Diagnosis.

2. Improve the Model's Capacity to Predict and Classify 14 Different Diseases (Multi-classification) in Chest X-rays.

3. Utilize a Comprehensive and High-Quality Dataset for Training and Testing the Deep Learning Model, Comprising Diverse and Representative Samples of Chest X-ray Images.

4. Fine-tune and Experiment with Different Hyperparameters to Optimize Model Performance (accuracy).

**Methodology:**

Begin:

Preprocessing: Load the data and resize the data

Split the data into training and validation set

Set up data loaders for training and validation set

Define the loss function

Define the model architecture:

Define the model architecture

Train the model:

Define hyperparameters for the model and train and evaluate the model

Evaluate the model:

Evaluate the model on testing dataset

Output the results

End

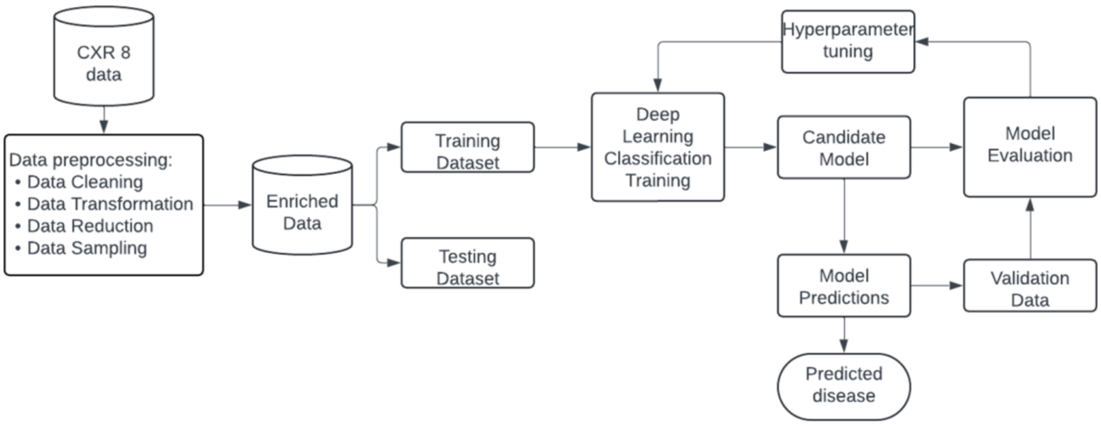


Fig: Overall flow

Extraction of dataset and preprocessing:

The image from datasets are loaded and are reduced in dimensions from 1024 X 1024 to 224 X 224 which helps improve the performance of a model during training

Split the data into training and validation set:

The dataset which is loaded consists of 1,12,120 images which is then divided into two sets out of which 86,524 images are used for training and validation purposes and 25,596 images are used for testing purposes. Further 86,524 images are divide into training and validation datasets in which 80% of images are used for training purposes and remaining 20% for validation purposes

Define the loss function:

A loss function is used to assess how well a model's predictions align with the actual target values (labels) in the training data. The primary purpose of a loss function is to measure the difference between predicted and actual values, providing a single scalar value that represents the "loss" or "error" of the model's predictions. In our case we have used a focal loss function as it is one of the majorily and best loss function and give very accurate results. It focus more on hard to classify example than well classified example.

$$\text{Focal Loss}(pt) = -\alpha t(1-pt)\gamma \log(pt)$$

*pt* is the model's estimated probability for the true class.
*αt* is a weighting factor for the class, used to balance the importance of positive/negative examples.
$\gamma$ is a focusing parameter that adjusts the rate at which easy examples are down-weighted.

Train the model:

After defining the configurations of each model we will then train each model on train datasets.

Evaluate the model:

After the training is completed the model is evaluated on the test dataset which consists of 25,596 images and a roc curve is plotted to know the accuracy of each class. A ROC(Receiver operating characteristic) curve is graph that shows how a binary classification model performs at different threshold values.

Output the results:

Once the model is trained and tested on the datasets model is feed with some random images and output is generated which gives the probability of a particular class

**Results and Conclusions:**

| Models | Layers | Eochs | Learning Rate | Parameter s | Validation Accuracy | Testing Accuracy |
|---|---|---|---|---|---|---|
| Denseset1 21 | 121 | 25 | 1e-04 | 7.97 million | 68.5% | 66% |
| VIT | 12 | 50 | 1e-05 | 86 million | 84.2 | 65% |

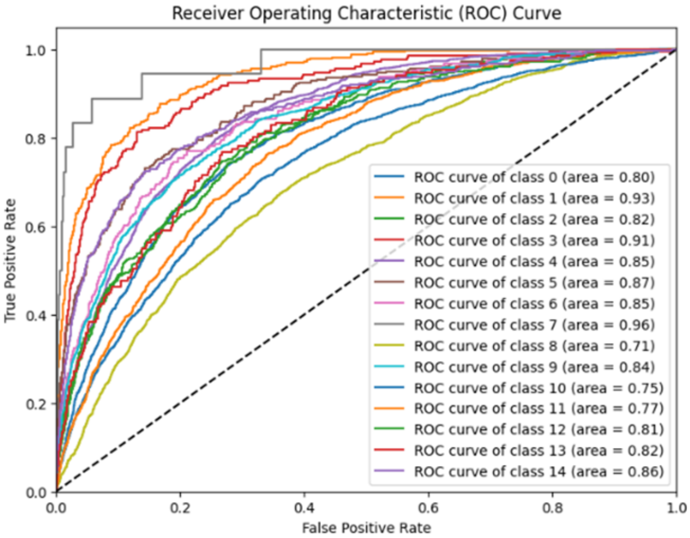


Fig: Validation ROC Curve for VIT

The curve shown in the fig is a validation roc curve for VIT model. We can see from the curve that for most of the classes AUC is almost 0.8 which is less than resnet50 but is still good

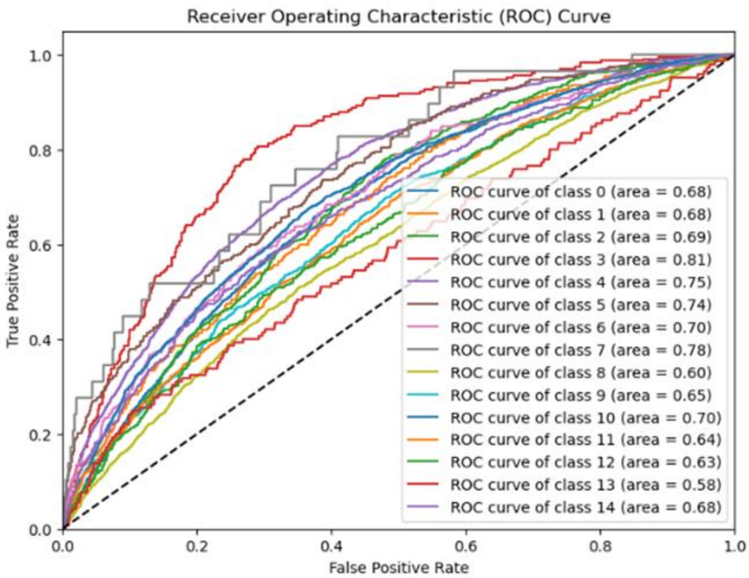


Fig: Validation ROC curve for densenet121

The curve shown in the fig is a validation roc curve for densenet121. We can see from the curve that for most of the classes AUC is between 0.6 to 0.8 which is less as compared to both resnet50 and VIT.

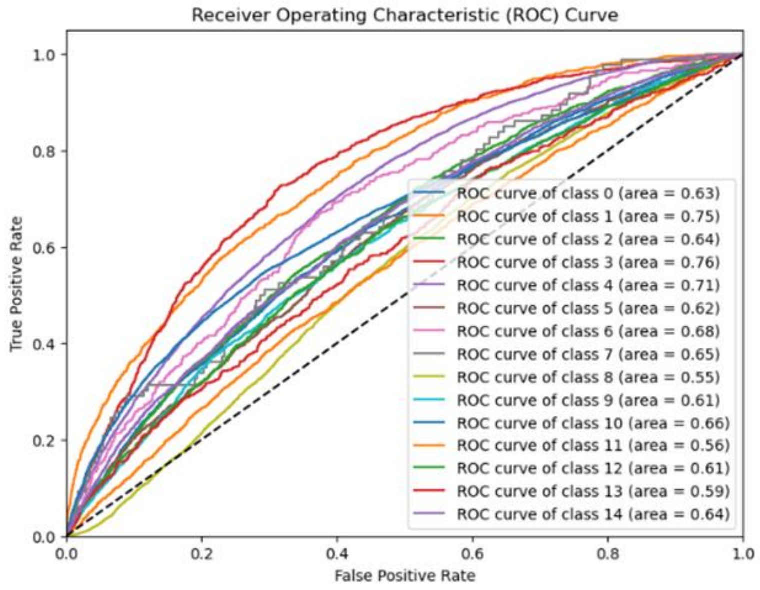


Fig: Testing ROC Curve for VIT

The curve shown in the fig is a testing roc curve for VIT. We can see from the curve that for most of the classes AUC is between 0.6 to 0.8 which indicates that the model is not working more accurately as compared to resnet50.

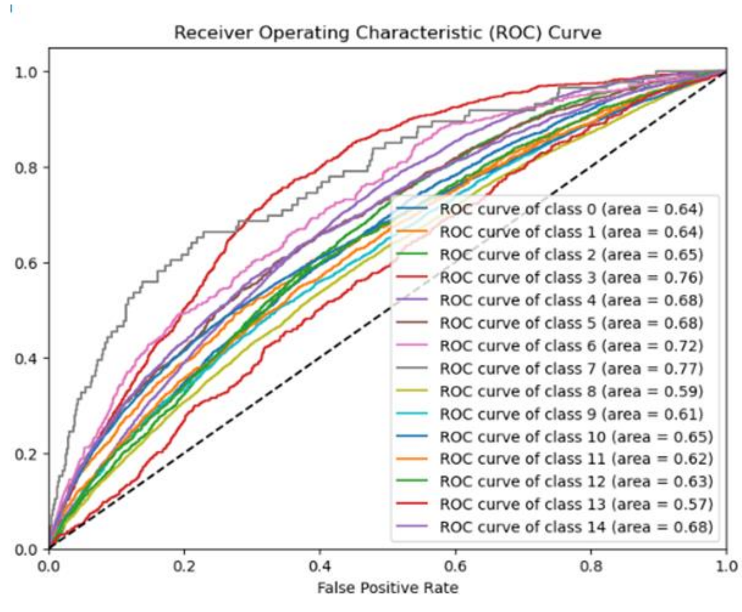


Fig: Testing ROC Curve for densenet121

The curve shown in the fig is a validation roc curve for densenet121. We can see from the curve that for most of the classes AUC is between 0.6 to 0.8 which is less as compared to resnet50.

In a comparative analysis of convolutional neural network architectures DenseNet121 and Vision Transformers we can see that Densenet121 outperforms VIT in terms of testing accuracy.

- This could be due to the fact that Densenet 121 architecture consists of direct connections with all the previous layers leading to concatenation of outputs therefore addressing the problem of vanishing gradient descent with direct gradient flow and the dense blocks tend to learn more accurately from a complex dataset.
- On the other hand, Vision Transformers have limitations affecting their performance compared to Densenet121. Vision Transformers rely heavily on large datasets for learning spatial dependencies through self-attention mechanisms, limiting their applicability in data-scarce scenarios.

**Scope for future work:**

- Improved accuracy of the deep learning model: In the ongoing development of machine learning models for medical applications, improving accuracy is a critical aspect of future scope. This focus on accuracy is particularly important for models diagnosing diseases from chest X-ray images.
- Integration with Clinical Systems: Integration of the developed deep learning model with existing clinical systems to facilitate real-time analysis and diagnosis, aiding healthcare professionals in making informed decisions.
- Collaboration with Healthcare Institutions: Collaboration with healthcare institutions for clinical validation and deployment of the developed deep learning model in real-world settings, ensuring its effectiveness and usability in a clinical environment.
- Development of a User-Friendly Interface: Designing a user-friendly interface for the deep learning model, making it accessible to healthcare professionals with varying levels of expertise in medical imaging interpretation.
- Research on Explainability and Interpretability: Conducting research on explainability and interpretability of the deep learning model to provide insights into its decision-making process, enhancing trust and acceptance among healthcare professionals.