Deep Learning for Pulmonary Diseases Detection Using Chest X-Ray

Aditya Ingole[1], Yuvraj Patil[1], Yashraj Wawkar[1], Aboli Deole[2]

[1]Student, Dept. of Artificial Intelligence and Machine Learning, PES's Modern College of Engineering, Pune, Maharashtra, India
[2]Assistant Professor, Dept. of Artificial Intelligence and Machine Learning, PES's Modern College of Engineering, Pune-05, Maharashtra, India.

Abstract - Pulmonary illnesses pose an enormous healthcare challenge globally, necessitating accurate and well-timed prognoses for effective remedies. Deep knowledge of Pulmonary ailment Detection using Chest X-rays provides a progressive strategy to decorate diagnostic accuracy and performance. Leveraging deep neural networks and a carefully curated dataset of chest X-ray photos, this assignment aims to automate the identification of pulmonary illnesses consisting of pneumonia, tuberculosis, emphysema and many more. The deep mastering version, educated and first-rate-tuned on this dataset, offers the potential to not only most effectively detect illnesses with high precision but additionally help healthcare specialists in early diagnosis, in the end enhancing patient results. The challenge's multifaceted technique consists of records preprocessing, model choice and training, interpretability, deployment in a scientific place, and non-stop collaboration with medical examiners to ensure both technological robustness and ethical compliance. As pulmonary disorder detection and healthcare technologies hold to adapt, this mission stands at the leading edge of innovation, presenting a promising method to increase the abilities of healthcare practitioners and deliver extra timely and accurate diagnoses.

Keywords: Pulmonary diseases, healthcare, diagnosis, deep learning, chest X-ray, diagnostic accuracy, efficiency, deep neural networks

1. INTRODUCTION

Nearly 545 million individuals currently live with a chronic respiratory condition, representing 7·4% of the world's population, which provides additional evidence of the large health contribution of chronic respiratory diseases to premature morbidity and mortality. Consequently, there has been substantial growth in the field of automated medical image classification.

This endeavour seeks to categorise medical images into specified groups. Lately, Deep Learning (DL) has emerged as a prevalent and extensively applied technique for creating medical image classification tasks.

Further, DL models produced more effective performance than traditional techniques using chest X-ray images from patients suffering from pulmonary diseases.

The DL architectures illustrated effective predictive ability. On chest X-ray images, multiple tasks were performed on DL models, including tuberculosis identification, tuberculosis segmentation, large-scale recognition, and Radio-graph classification. The automated classification of chest X-ray images using DL models is growing rapidly and choosing an appropriate region of interest (ROI) on chest X-ray images was used to treat. Furthermore, applying the DL modes helps to avoid problems that take a long time to solve in traditional approaches. However, these models require large volumes of well-labelled training samples. The DL architectures illustrated effective predictive ability. On chest X-ray images, multiple tasks can be performed on DL models, including tuberculosis identification, tuberculosis segmentation, atelectasis, pneumonia, silico-tuberculosis, fibrosis, Emphysema Radiograph classification and many more. Many researchers have done investigations to relate machine learning schemes for prediction of X-ray image diagnostic information. With the control of computers along with the huge volume of records being unrestricted to the public, this is high time to resolve this complication. This solution can put up decreasing medical costs with the enlargement of computer science for health and medical science projects. For the implementation, the NIH chest X-ray image dataset is collected from Kaggle repository, and it is fully an open-source platform. A new hybrid algorithm is introduced in this paper and this algorithm is successfully applied on the above-mentioned dataset to classify lung disease. The main contribution of this research is the development of this new hybrid deep learning algorithm suitable for predicting lung disease from X-ray images.

2. OBJECTIVES

- The primary goal of this project is to develop a robust system that can detect and analyze different lung diseases using chest x-ray images. Traditional methods can make it difficult to diagnose lung diseases such as pneumonia, tuberculosis, lung cancer, and COPD. The model is developed with sophisticated machine learning and deep learning methods to detect unusual patterns and abnormalities in chest x-rays, revealing different lung illnesses. This system aims to improve diagnostic precision, minimize human mistakes, and speed up the diagnostic procedure, ultimately leading to improved patient results.
- This model is valuable for both diagnosing patients and conducting statistical analysis in hospitals and medical research facilities. Researchers can use the model to collect and analyze extensive collections of chest x-ray images, making it easier to conduct epidemiological studies and clinical trials. The model helps medical research by offering information on disease frequency, advancement, and response to treatment to aid in decision-making. This thorough examination helps to identify trends, comprehend patterns of diseases, and create new treatment guidelines, ultimately pushing forward the field of pulmonary medicine.
- This invention was created to aid physicians and radiologists in their routine medical duties. Doctors can improve their clinical assessments by incorporating an extra perspective from automated analysis, boosting their certainty when diagnosing patients. For radiologists, it acts as a quick initial assessment tool, identifying possible areas of concern that need more detailed examination. The model effectively understands and analyzes extensive x-ray images, leading to a notable decrease in the workload of healthcare professionals. This enables them to concentrate on intricate cases and provide personal attention to patients. In the end, the goal of this technology is to enhance the efficiency and effectiveness of diagnosing and treating lung diseases, ultimately resulting in improved outcomes.

3. LITERATURE REVIEW

Image classification involves assigning a predetermined category label to an image and is a component of computer vision. CNNs and DNNs have revolutionized the process of image classification by having the capability to extract features from hierarchical image data. EfficientNet-B2 and Inception V3 are two well-known CNN structures that have demonstrated outstanding performance in different image classification assignments. Image manipulation in deep CNNs includes a variety of methods and procedures that ready images for efficient analysis and feature extraction. This procedure is important for various uses, such as medical imaging, where precise and thorough analysis of images is necessary. CNNs such as EfficientNet-B2 and Inception V3 are crucial for improving the effectiveness of Image processing tasks in the identification of pulmonary diseases through chest X-rays (CXRs) analysis. Inception V3, is an advanced version of the Inception architecture, it was developed by et al Szegedy. It uses various methods to improve network performance, such as factorized convolutions, batch normalization, and auxiliary classifiers. Inception V3 is optimized to effectively process high-res images, making it appropriate for medical imaging uses.

3.1 Main characteristics

Factorized Convolutions: Converts convolutions into smaller operations for decreased computational complexity. Auxiliary classifiers are additional classifiers that offer extra gradient signals in the training process, enhancing convergence and reducing the vanishing gradient issue. Normalization of each layer is implemented using Batch Normalization to enhance and speed up the training process. Design of buildings and structures. The structure of Inception V3 includes several Inception modules, which have concurrent convolutions with various kernel sizes. This capability of extracting features at multiple scales enables the network to capture a variety of features, improving its capacity to classify intricate images.

3.2 EfficientNet-B2:

Overview

EfficientNet-B2, introduced by Tan and Le, is a subset/subtype of the EfficientNet family, which uses a compound scaling method to balance network depth, width, and resolution. This systematic scaling approach results in models that achieve high accuracy while maintaining computational efficiency.

Key Features

- Compound Scaling: Simultaneously scales depth, width, and resolution using a set of predefined coefficients, optimizing the model for performance and efficiency.
- MBConv Blocks: Mobile Inverted Bottleneck Convolution blocks with squeeze-and-excitation (SE) optimization, enhancing feature representation and reducing the number of parameters.
- Swish Activation Function: This activation function improves training stability and model performance compared to the traditional ReLU function.

Architecture

EfficientNet-B2 employs MBConv blocks, which include complex depth-wise separable convolutions and squeeze-and-excitation components. These blocks help in capturing fine-grained features while maintaining computational efficiency. The compound scaling method ensures that the network is efficiently balanced in terms of its capacity to learn from data without overfitting.

3.3 InceptionV3

Inception V3, EfficientNet-B2, VGG16, VGG19, and ResNet-50 showcase a range of convolutional neural network (CNN) designs, each possessing unique qualities and features. Inception V3 is renowned for its complex structure and use of inception modules, making it highly effective in capturing features at various levels, thus being well-suited for tasks demanding detailed feature identification like image classification and object detection. EfficientNet-B2, from the EfficientNet family known for its efficient scaling, finds a middle ground between model size and performance, making it appropriate for environments with limited resources while still maintaining accuracy.In contrast, VGG16 and VGG19, although less complex in design than the rest, are famous for being easy to comprehend and put into practice. Despite their deep structure, they are relatively straightforward, consisting of sequential layers with small convolutional filters. These models are often used as baseline architectures or for transfer learning due to their simplicity and effectiveness. Meanwhile, ResNet-50 introduces the concept of residual learning, which addresses the vanishing gradient problem in deep networks, enabling the training of significantly deeper architectures. This results in improved performance and convergence speed compared to earlier models. ResNet-50 is widely adopted in various computer vision tasks, especially when depth and accuracy are paramount.

3.4 Challenges

- The main challenges faced by conventional methods of Pneumonia detection such as chest radiographs is that they are subject to inter-class variability and ultimately the diagnosis is dependent on the expertise of the clinician in detecting early traces of pneumonia.
- One limitation of Emphysema detection is the lack of reference standard for the emphysema label based on CT imaging or confirmed emphysema diagnosis. The models presented here are trained only to emulate the performance of a radiologist identifying emphysema on a chest X-Ray and are evaluated in that context also. This does not provide any indication of how well CXR-based analysis compares with more accurate reference standards such as quantitative CT and/or clinical diagnosis.
- The authors acknowledge some limitations of their study on the model of Lung Cancer detection system, such as the use of a single hospital dataset, the potential increase of false positives in a screening cohort, the need for an observer's performance study and the exclusion of benign nodules/masses from their model.
- The main gap in the research paper and module on CF-ILDs is the dataset they used to train and test the model was collected from a single source, which limits the generalization of the algorithm. The other gap was that the algorithm was not validated on other types of ILDs. The algorithm did not provide any information on the location, extent or severity of the disease, and it did not incorporate any clinical or laboratory data which may improve the accuracy and confidence of the CF-ILDs detection model.

3.5 Integrating Inception V3 and EfficientNet-B2

Inception V3, created by Google, is a sophisticated Convolutional Neural Network (CNN) structure specially made to effectively capture complex features across various scales. This approach at multiple scales enables the network to gather varied and abundant features from input images. Additionally, Inception V3 employs factorized convolutions, where larger convolutions are split into smaller, sequential operations (e.g., a 3x3 convolution is decomposed into a 1x3 followed by a 3x1 convolution). This factorization reduces the computational cost while maintaining performance. Another important feature is the use of auxiliary classifiers during training, which provide intermediate gradient signals to improve convergence and training stability using batch normalization extensively aids in standardizing the inputs for every layer, which in turn decreases internal covariate shift and speeds up the training process. Dropout and label smoothing are regularization methods used to avoid overfitting and enhance the model's ability to generalize. The combination of these features makes Inception V3 very successful in achieving precise results in various image classification assignments, such as intricate tasks like medical imaging. EfficientNet-B2 belongs to the EfficientNet series, praised for its outstanding blend of performance and computational efficiency. The hallmark of EfficientNet-B2 is its use of compound scaling, a method that uniformly scales the network's depth (number of layers), width (number of channels), and resolution (input image size) using a compound coefficient. This balanced scaling ensures optimal performance while minimizing computational cost. The architecture employs Mobile Inverted Bottleneck Convolutions (MBConv) as its core building blocks, which combine depth wise separable convolutions and inverted residuals. These MBConv blocks significantly reduce the number of parameters and computational load without compromising the ability to learn complex features. Additionally, EfficientNet-B2 integrates Squeeze-and-Excitation (SE) optimization, where SE blocks recalibrate channel-wise feature responses adaptively by modelling interdependencies between channels. This enhances the network's representational power. The use of the Swish activation function further improves training stability and performance. Overall, EfficientNet-B2 achieves high accuracy with fewer resources, making it an excellent choice for applications that require both performance and efficiency, such as deployment in resource-constrained environments or transfer learning tasks.

The performance of image classification tasks greatly improves with the implementation of convolutional neural networks (CNNs) utilizing important techniques like Pre-training and fine-tuning in the realm of deep learning. Pre-training consists of training a neural network on a flexible, universal dataset like ImageNet to acquire a diverse set of features that can be applied to different tasks. This process creates a robust set of initial weights that capture fundamental visual patterns. After being pre-trained, the model can be adjusted on a smaller dataset specifically for the task. Fine-tuning the model requires swapping out the last layers of the pre-trained model to align with the attributes in the new dataset, followed by training the entire network on the new data. This approach leverages the general features learned during pre-training while allowing the model to adapt to the specific characteristics of the new task.

During fine-tuning, early layers of the network, which capture generic features like edges and textures, are often frozen to retain their pre-trained weights. The later layers, which capture more task-specific features, are updated to learn the new data's intricacies. This gradual unfreezing, combined with a reduced learning rate, helps in refining the model without overfitting. Optimizers play a critical role in both pre-training and fine-tuning stages. The Adam optimizer is commonly used due to its adaptive learning rate capabilities, which make it effective for a wide range of tasks. Adam combines the advantages of both AdaGrad and RMSProp, adjusting the learning rate for each parameter based on the first and second moments of the gradients, which accelerates convergence and improves performance. Additionally, learning rate scheduling techniques, such as step decay or reduction on the plateau, are employed to fine-tune the learning rate during training, ensuring stable and efficient convergence. These combined strategies of pre-training, fine-tuning, and adaptive optimization enable deep learning models to achieve high accuracy and generalization on specific tasks with limited data.

CNNs have transformed the field of image classification by showcasing incredible abilities in autonomously learning and identifying patterns and features within images. CNNs can capture hierarchical features, from basic edges and textures to intricate shapes and object parts, due to their layered architecture that imitates the human visual system. CNNs are very effective for various image classification tasks, such as identifying handwritten digits and complex objects in high-resolution images, due to their capacity to learn hierarchical representations. Cutting-edge designs like Inception V3, EfficientNet, and ResNet have elevated performance levels by integrating new design aspects such as multi-scale feature extraction, effective utilization of computational resources, and enhanced gradient flow via residual connections. These models have been trained in advance on extensive datasets such as ImageNet, allowing them to perform effectively on a wide range of datasets and tasks. Moreover, by employing transfer learning and fine-tuning, these models can be customized for specific tasks even with a scarcity of labelled data, leading to a remarkable decrease in training time and computational resources. Optimization methods like Adam, along with practices such as batch normalization and dropout, improve the training procedure by promoting quicker convergence and increasing resistance to overfitting. The result is highly accurate and reliable image classification systems that are now integral to numerous applications, including medical diagnosis, autonomous driving, and real-time video analysis.

3.5.1 Advantages of Inception V3 and EfficientNet-B2

Inception V3 and EfficientNet-B2 are two sophisticated convolutional neural network designs that provide notable benefits in tasks related to image classification. Inception V3 enables us to effectively identify both fine and coarse details in images, making it highly accurate in diverse classification tasks. Additionally, the model employs factorized convolutions to reduce computational costs and auxiliary classifiers to stabilize training and enhance convergence speed. Batch normalization and regularization methods like dropout and label smoothing enhance its performance by avoiding overfitting and ensuring seamless training. EfficientNet-B2 is known for its combination of efficiency and performance, which is achieved through a compound scaling strategy. This method scales depth, width, and resolution in a coordinated manner, optimizing the model for both accuracy and resource usage. The core of EfficientNet-B2 consists of Mobile Inverted Bottleneck Convolutions (MBConv) with depth wise separable convolutions and squeeze-and-excitation (SE) optimization, which enhance feature extraction while keeping the parameter count low. The use of the Swish activation function contributes to training stability and overall performance. EfficientNet-B2's design is very efficient, allowing it to be used in environments with limited resources while still achieving top accuracy levels. Both models are great choices for transfer learning, enabling them to be efficiently customized for tasks with minimal data, highlighting their flexibility and strength in image classification tasks.

Inception V3 operates by utilizing a complex yet efficient architecture designed to capture rich and diverse features at multiple scales.This method with multiple branches helps the network to detect fine-grained and coarse features at the same time. The module then concatenates these parallel outputs, creating a comprehensive feature map that represents various aspects of the image. To enhance computational efficiency, Inception V3 employs factorized convolutions, which decompose larger convolutions into smaller, sequential operations, significantly reducing the model's computational burden without sacrificing accuracy.Moreover, Inception V3 incorporates extra classifiers while training to give intermediate gradient signals that help stabilize the training process and speed up convergence. Batch normalization is widely applied to standardize inputs for every layer, enhancing both the speed of training and the overall performance of the model. EfficientNet-B2, a member of the EfficientNet group, utilizes compound scaling to uniformly adjust the network's depth, width, and resolution by employing a compound coefficient. This method ensures a balanced increase in model size and complexity, leading to optimal performance and efficiency. The architecture of EfficientNet-B2 is built around Mobile Inverted Bottleneck Convolution (MBConv) blocks, which include depthwise separable convolutions and inverted residuals. These blocks reduce the number of parameters and computational load while maintaining high feature extraction capabilities. Furthermore, EfficientNet-B2 incorporates Squeeze-and-Excitation (SE) optimization, where SE blocks adaptively recalibrate channel-wise feature responses, enhancing the network's ability to capture essential features. The Swish activation function used in EfficientNet-B2 improves model performance by allowing smoother gradient flow during training. This combination of advanced architectural elements enables EfficientNet-B2 to achieve high accuracy with reduced computational resources, making it an efficient and powerful model for image classification tasks.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 09 | Sep 2024 www.irjet.net p-ISSN: 2395-0072

4. Methodology

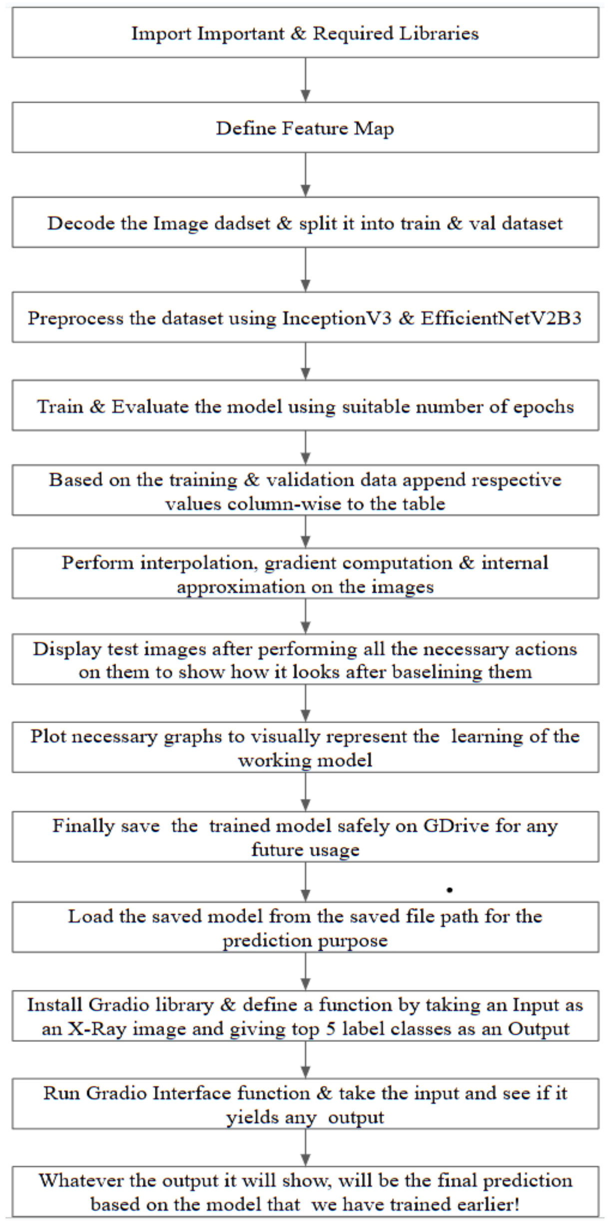

Fig. 1 Dataflow Diagram

4.1 Data Collection

The dataset utilized for this study contains chest X-ray images extracted from a publicly available repository i.e. NIH Chest X-ray Dataset. This dataset is available on kaggle and provides a collection of labelled chest X-ray images, including normal lung samples and the ones in which the disease is detected. The images are preprocessed to ensure consistency in size and resolution, resized to 150x150 pixels for EfficientNet-B2 and for Inception V3.

4.2 Preprocessing

Preprocessing steps are significant for preparing the X-ray images as input into the neural networks. The following steps are followed below:

- Normalization: Pixel values are normalized in the range of [0, 1] to ensure consistency of the input data.
- Data Augmentation: To enhance the efficiency and performance of the models and solve the problem of low scalability various data augmentation techniques are applied, also we use an image generator for this purpose. These variations help in simulating real-world variations and prevent overfitting.

4.3 Model Architecture

- EfficientNet-B2

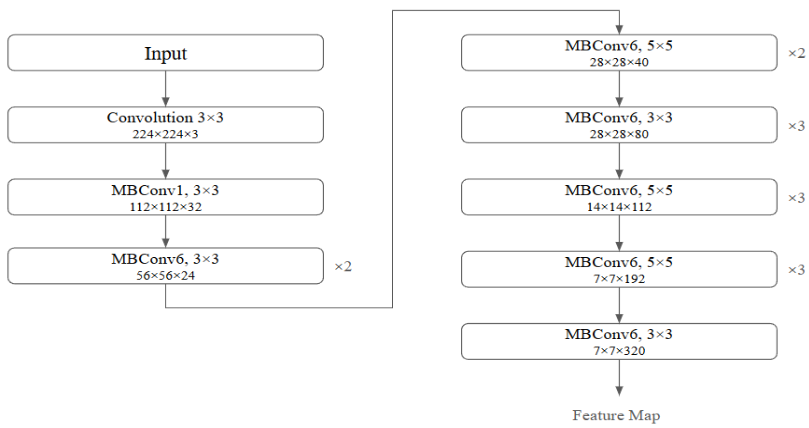

Fig. 2 EfficientNetB2 Architecture

EfficientNet-B2, a subtype of the EfficientNet family, uses compound scaling to balance network depth, width, and resolution in the neural networks. The architecture includes:

MBConv Blocks: Mobile Inverted Bottleneck Convolution layers with depth wise separable convolutions and inverted residuals to reduce parameters and computational load.

Squeeze-and-Excitation (SE) Optimization: SE blocks recalibrate channel-wise feature responses to improve representational power.

Swish Activation Function: Enhances training stability and model performance.

- Inception V3

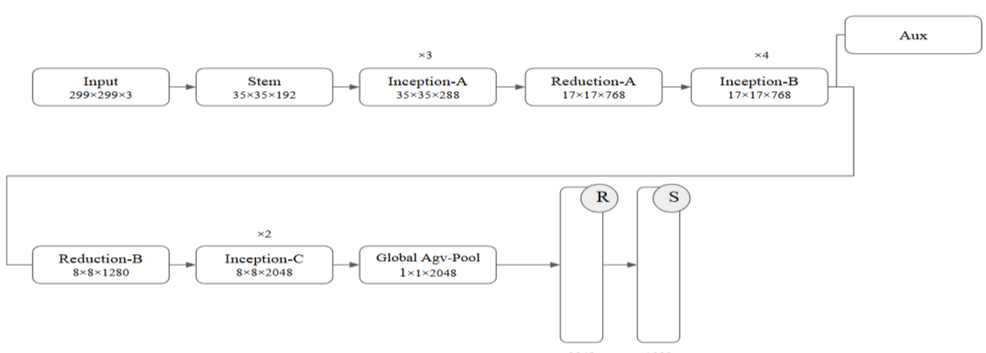

Fig. 3 InceptionV3 Architecture

Inception V3 employs a sophisticated architecture designed for multi-scale feature extraction:

- Inception Modules: Parallel convolutions with various filter sizes (1x1, 3x3, 5x5) and pooling operations to capture diverse features.
- Factorized Convolutions: Decomposition of larger convolutions into smaller ones to reduce computational costs.
- Auxiliary Classifiers: Intermediate classifiers that provide additional gradient signals during training, aiding in convergence and stability.
- Batch Normalization: Applied throughout the network to normalize inputs to each layer, reducing internal covariate shift and accelerating training.

4.4 Pre-training and Fine-tuning

Both models are pre-trained on the ImageNet dataset to leverage the learned feature representations. Pre-training involves training EfficientNet-B2 and Inception V3 on ImageNet, a massive dataset containing millions of labeled images in thousands of categories. Fine-tuning involves adjusting the pre-trained models using data from the NIH Chest X-ray dataset. (e.g., normal, pneumonia and other pulmonary diseases). Early layers are initially frozen to retain general feature representations, while later layers are updated to learn task-specific features. Gradually, more layers are unfrozen for complete fine-tuning.

4.5 Training Procedure

Data Splitting: The dataset is divided into training, validation, and test sets with an 80-10-10 distribution. The Adam optimizer is utilized for its ability to adjust the learning rate adaptively. Initial learning rates are set low and adjusted using learning rate schedulers.

- Loss Function: Cross-entropy loss is employed for multi-class classification.
- Batch Size and Epochs: Batch sizes of 32 and 50 epochs are set for training, with early stopping based on validation performance to prevent overfitting.

4.6 Evaluation Metrics

Model performance is evaluated using standard metrics:

- Accuracy: Proportion of correctly classified instances. Precision, Recall, and F1-Score: These are calculated class-wise to measure how well our model can detect each of the conditions.

4.7 Implementation

The models are implemented using deep learning frameworks such as TensorFlow and Keras. Training and evaluation are conducted on GPU-accelerated environments to expedite the computational processes. The trained models are then evaluated on the test set to determine their final performance metrics.

4.8 Result Analysis

Results are analyzed to compare the performance of EfficientNet-B2 and Inception V3 in detecting pulmonary diseases. Key insights and comparative analysis are provided, highlighting strengths and potential areas for improvement in each model. This methodology outlines a comprehensive approach to leveraging EfficientNet-B2 and Inception V3 for the detection of pulmonary diseases using chest X-ray images, ensuring robust and accurate model development through pre-training, fine-tuning, and detailed evaluation.

5. Implementation

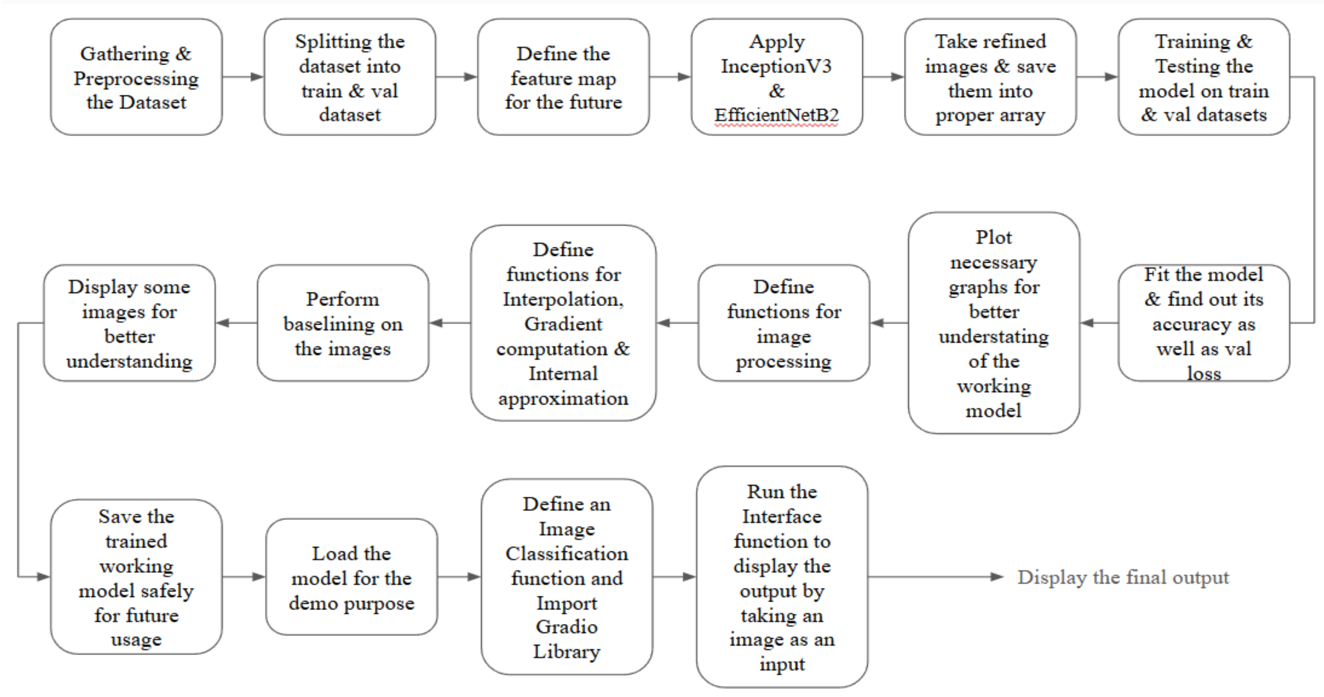

Fig. 4 UML Diagram

5.1 Image Output with Disease Classification and Image Segmentation

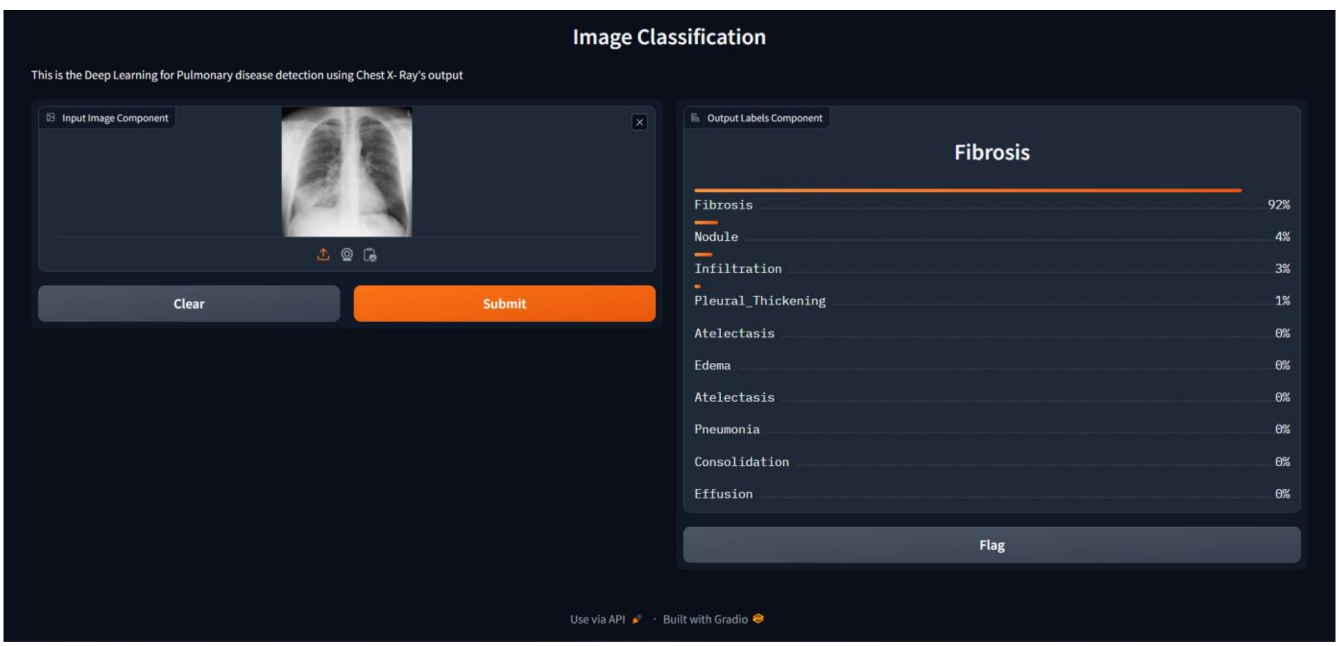

Fig. 8 Output: Prediction

To provide visual interpretability of the model's predictions, integrated gradients are calculated for the input image. Integrated gradients highlight the important pixels that contribute to the model's decision, which helps in understanding the model's focus areas. The integrated gradients are normalized and visualized alongside the original image, showing how different parts of the X-ray contributed to the classification made by EfficientNet-B2 and Inception V3. This step enhances transparency and trust in the deep learning model's decision-making process.

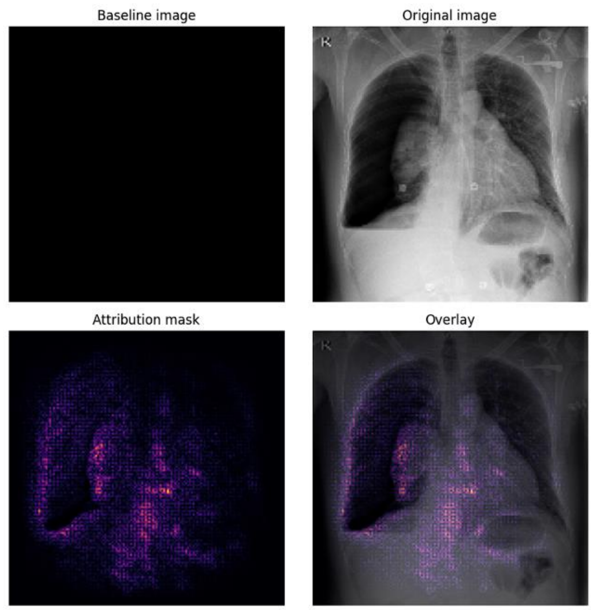

Fig.9 Lung Segmentation & Overlay

6. CONCLUSION

In this article, our goal is to propose a deep learning-based approach to classify pulmonary diseases from chest X-ray images using CNN based algorithms such as the EfficientNet B2 and Inception V3. In this framework, we pretrained the architecture using the ImageNet dataset. For Model training we used the NIH (National Institute of Health) Chest X-ray dataset.

While most pulmonary diseases are commonly confirmed by a single doctor, allowing for the possibility of error, our model can be regarded as a two-way confirmation system. Our results suggest that our CNN based models can be used to improve diagnosis relative to traditional methods, which may improve the quality of treatment and reduce the time consumed. When compared with the previous methods, our approach can effectively detect the inflammatory region.

7. REFERENCES

[1] Patrik Szepesi, László Szilágyi, "Detection of pneumonia using convolutional neural networks and deep learning", Biocybernetics and Biomedical Engineering, Volume 42, Issue 3,2022, Pages 1012-1022, ISSN 0208-5216, https://doi.org/10.1016/j.bbe.2022.08.001.

(https://www.sciencedirect.com/science/article/pii/S0208521622000742)

[2] Çallı E, Murphy K, Scholten ET, Schalekamp S, van Ginneken B (2022) Explainable emphysema detection on chest radiographs with deep learning. PLoS ONE 17(7): e0267539. https://doi.org/10.1371/journal.pone.0267539

[3] Shimazaki, A., Ueda, D., Choppin, A. et al. Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method. Sci Rep 12, 727 (2022). https://doi.org/10.1038/s41598-021-04667-w

[4] H. Nishikiori, K. Kuronuma, K. Hirota, et al., "Deep learning algorithm to detect fibrosing interstitial lung disease on chest radiographs," Eur. Respir. J., vol. 4, no. 1, Art. no. e02269-2021, Sep. 2022, doi 10.1183/13993003.02269-2021. https://erj.ersjournals.com/content/early/2022/09/30/13993003.02269-2021

[5] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE Conference on computer Vision and pattern recognition (CVPR); 2017. p. 3462–71. https://doi.org/ 10.1109/CVPR.2017.369.

[6] Gu Y, Lu X, Yang L, Zhang B, Yu D, Zhao Y, Gao L, Wu L,.Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs - ScienceDirecthttps://www.sciencedirect.com/science/article/abs/pii/S001048251830310X

[7] Setio AAA, Traverso A, de Bel T, Berens MSN, van den Bogaard C, Cerello P, Chen H, Dou Q, Fantacci ME, Geurts B, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge - ScienceDirecthttps://www.sciencedirect.com/science/article/abs/pii/S1361841517301020

[8] Zhu W, Liu C, Fan W, Xie X DeepLung. Deep 3D dual path nets for automated pulmonary nodule detection and classification. In: Proceedings of the IEEE winter conference on applications of computer vision (WACV); 12–15 March 2018. p. 673–81. Lake Tahoe, NV, USA. https://ieeexplore.ieee.org/document/8354183

[9] Kong W, et al. YOLOv3-DPFIN: a dual-path feature fusion neural network for robust real-time sonar target detection. IEEE Sensor J 1 April 1, 2020;20(7): 3745–56. https://doi.org/10.1109/JSEN.2019.2960796.

[10] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, vol. 9351. Berlin/Heidelberg, Germany: Springer; 2015. p. 234–41. http://refhub.elsevier.com/S2352-9148(20)30029-0/sref15

[11] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

keywords: {Computer architecture; Convolutional codes; Sparse matrices; Neural networks; Visualization; Object detection; Computer vision}https://ieeexplore.ieee.org/xpl/conhome/7293313/proceeding