Deep Learning-Based Pediatric Pneumonia Detection in Children: A Comparative Analysis of Techniques for Improved Diagnosis

Prabhav Ravi Tammanashastri
Dept. of Computer Science and Engineering
Amrita School of Computing,Bengaluru
Amrita Vishwa Vidyapeetam,India
BL.EN.U4CSE21162

Dumpa Venkata Sai Jagadeesh
Dept. of Computer Science and Engineering
Amrita School of Computing,Bengaluru
Amrita Vishwa Vidyapeetam,India
BL.EN.U4CSE21052

Suhas S Bhat
Dept. of Computer Science and Engineering
Amrita School of Computing,Bengaluru
Amrita Vishwa Vidyapeetam,India
BL.EN.U4CSE21200

Dr Rimjhim Singh*
Dept. of Computer Science and Engineering
Amrita School of Computing,Bengaluru
Amrita Vishwa Vidyapeetam,India

I. Introduction

Pediatric pneumonia remains a significant cause of morbidity and mortality among children globally.Treatment options for pneumonia are based on the patient's age, contraction, and other medical issues. Thus, the keys to lowering mortality are early intervention and accurate diagnosis. The increasing number of medical equipment and digital recording systems has generated a vast volume of data, which deep learning has been using to diagnose pathologies including glaucoma, brain tumours, and skin cancer. These results are encouraging.

We use Transfer learning Models to allow the utilization of pre-trained models on large datasets for tasks with limited training data including computer vision, natural language processing, and healthcare.One well-known example is the Visual Geometry Group (VGG), which provides a number of versions, including VGG16 and VGG19, which are well-known for being straightforward and efficient when it comes to feature extraction tasks. The ability of ResNet (Residual Network) models to reduce the vanishing gradient issue and produce more reliable representations has also contributed to their growing popularity. These models have deep architectures with skip connections. Inception models improve their performance in image recognition tasks, such as InceptionV3 and InceptionResNetV2. Compound scaling is a recent development in EfficientNet, which achieves state-of-the-art performance with enhanced computational efficiency by balancing model depth, width, and resolution. Healthcare applications can also benefit from domain-specific transfer learning models like MIMIC-CXR for radiological imaging tasks and CheXNet for chest X-ray processing.

Furthermore, techniques like transfer learning enable CNNs to leverage pre-trained models, reducing the need for extensive training data and computational resources. Through the integration of these components, CNNs play a crucial role in enhancing diagnostic accuracy and efficiency in pediatric pneumonia detection, ultimately improving patient outcomes and reducing healthcare burdens.

Add objectives of paper

II. Literature Survey

Spencer er al [1] paper on necrotizing pneumonia in children illuminates the evolving landscape of this condition, once considered uncommon but now on the rise. Current research underscores a diverse etiology, with pneumococcal infection, Methicillin-resistant Staphylococcus aureus, MRSA, and Panton-Valentine leucocidin PVL, staphylococcal infection playing pivotal roles. Clinical manifestations closely resemble uncomplicated pneumonia but exhibit heightened severity, demanding tailored interventions. The surveyed literature advocates for comprehensive management, emphasizing fluid balance, analgesia, and intravenous broad-spectrum antibiotics, the selection of which is guided by local resistance patterns. Despite the seriousness of necrotizing pneumonia, the literature consistently reports favorable outcomes, with a majority of children making a complete recovery. However, challenges exist, including a limited understanding of contributing mechanisms, variable clinical presentations, and the dependence on intensive care resources for severe cases. Further research is warranted to address these gaps and refine management protocols, ensuring optimal outcomes for pediatric patients with necrotizing pneumonia.

Chen et. al [2] have used some of the deep convolutional neural networks such as ResNet and DesNet. The method proposed by the authors enables feature operations, such as DLF and FLF. In the DLF module, we have two auxiliary classifiers CD, CR, and a fusion classifier CF. In the FLF, we have asymmetric networks that are based on DenseNet and ResNet using which we can form the feature extractor of Dual CheXNet. The authors have tried to investigate different configurations of FLF and DLF. It is denoted as DualCheXNet-1 and DualCheXNet-2. DualCheXNet-1 is with 50-layer ResNet and 121-layer DesNet. DualCheXNet-2 with 101-layer ResNet and 169 DesNet. The authors have compared the AUC scores and the corresponding single networks of six models on ChestXray-14.Out of 14 test sets, DualCheXNet-2 is performing better in predicting 11 Pathologies. For 14 Pathologies, the AUC score is 0.823.

Rutuja Pansare et al [3] tells that The dataset was used to train four different models. Transfer learning and convolutional neural networks formed the foundation of the models' development. CNN uses convolution layers, batch normalization, batching, Activation function as ReLu, and usage of Dropout and Dense layers for Training, validation, and Testing of the Images. To improve the learning rate of the model batch normalization is used. Max-pooling with a 2*2 pooling filter was applied after each convolutional layer in the basic CNN model. Every layer ended with ReLU, and the last layer with two nodes employed the sigmoid activation function. Dropout is a method to lessen the model's overfitting. Using dropout, a random selection of layer nodes is made to remain dormant for a certain amount of time. Transfer Learning uses the architecture of pre-trained models such as InceptionV3, VGG19, and VGG16 along with weights acquired during training on the ImageNet dataset. There were two approaches to the training. One method involves a previously trained model to maintain the initial weights and update the other weights to freeze the Convolutional Layers. In contrast, the model was being trained on the dataset. In the second model, when the model was trained, additional weights were added to the original ones. This block's output was further flattened before being fed to the layers that are connected fully, which applied activation, batch normalization, and dropout in the same way as the basic CNN model.

Liang et al [4] performs pediatric pneumonia diagnosis using deep residual networks which reveals a pioneering approach that amalgamates transfer learning, deep residual structures, and dilated convolution. The method efficiently tackles issues of overfitting and degradation through the integration of residual structures, enhancing the robustness of the diagnostic system. Dilated convolution ensures the preservation of crucial spatial information, offering sensitivity to object position and orientation in medical images. Transfer learning, initialized with parameters learned from extensive datasets, addresses challenges related to insufficient data and structured noise, optimizing model training. Despite these advantages, potential drawbacks include reliance on large-scale datasets, computational intensity associated with deep architecture, potential interpretability issues, and the necessity for thorough validation across diverse populations. In conclusion, the proposed method shows promise in advancing pediatric pneumonia diagnosis, but careful consideration of its dependencies and further research is essential for real-world applicability.

Habib et al [5] delves into a novel approach for pediatric pneumonia detection, employing an ensemble method based on CheXNet and VGG-19, two deep Convolutional Neural Networks (CNNs). Acknowledging the critical importance of timely diagnosis for patient recovery, the study leverages the feature extraction capabilities of VGG-19 and CheXNet to create a rich ensemble of features from X-ray photos. To address data irregularities, the ensemble incorporates a SMOTE, ROS, and RUS on the feature vector. Classification is performed using various ML techniques, with Random Forest exhibiting superior performance metrics on a standard dataset, achieving an impressive 98.93

Jain et.al [6] Several CNNs have been trained to distinguish between non-pneumonia and pneumonia in X-ray pictures. There are two convolutional layers in the first model and three in the second. The accuracy of the first two models is 85.27

Guan et. al [7] have worked on thorax disease with in Chest X-rays. The authors have proposed a CRAL (category-wise residual attention learning) framework for classifying chest X-ray images. The authors have attention mechanisms att1 and att2 in calculating the effectiveness of the CRAL framework. For the proposed model, when using backbone R- 50 the AUC for att1 is 0.8136 and for att2 is 0.8102. When using backbone D-121 the AUC for both att1 and att2 is 0.8157.

Alhudhaif et al[8] delve into recent advancements in the prediction of pneumonia for Covid19 patients through the CNN models on chest X-ray. Studies commonly employ transfer learning,utilizing architectures like DenseNet-201, ResNet-18, and squeezing in enhancing feature extraction and accelerate model development. A prevalent method for mitigating bias involves a stratified cross-validation approach, ensuring a balanced representation of pneumonia cases and COVID-19 across training and testing sets. Additionally, activation mapping techniques contribute to model interpretability by visually highlighting regions crucial to classification decisions. While these methods offer advantages such as improved accuracy, robust evaluation, and enhanced interpretability, considerations must be given to potential limitations, including adaptability of pre-trained models, computational overhead in cross-validation, and complexities in activation map interpretation. Overall, these findings showcase a diagnosis that is accurate for chest X-ray images in predicting Pneumonia for COVID-19 patients.

Wang et. al [9] have proposed a triple net model (A3) model for Computer Aided Diagnosis Task (CAD). For feature extraction, the model is going to use pre-trained Dense121as the backbone network. To find the relationship between the diagnosis-specific location of images, disease labels the diagnosis-specific feature channels the authors have used attention modules. The Triple Net model has been evaluated on chest X-ray images which have 112,120 images. Among the other deep-learning models the A3 has the highest average per class which is 0.826, classifying 13 out of 14 thoracic diseases.

Geraldo Braz et al [10] proposes a methodology suggests creating a new, customized network for the classification task by combining pre-trained network layers with an optimal search of topologies and hyperparameters. This results in the best possible convolutional neural network structure. The goal is to identify the network that best distinguishes between pneumonia caused by bacteria and viruses the best. To compare the approaches, we employ the F-Score. In this study, The base network, or backbone, is a pre-trained VGG16. In comparison to other pre-trained networks, the total number of layers and filters increases the number of weights. This network does not have time-consuming methods that eliminate feature representations and distribution. VGG16 is suited for general jobs in medical imaging, namely X-ray images, because of its behavior. During the optimization process, decisions are made on the number of filters, neurons, convolution, and fully connected blocks, among other parameters. One hundred iterations of optimization are carried out. After the network layers are frozen for about 20 epochs, all networks undergo 100 epochs of training with all layers to adjust specialized network weights. Using certain layers that are trained prior based on the VGG16 network is linked to a fully predicted CNN, The topology of CNN, and all of the hyperparameters for the task are estimated after the optimization process in order to maximize the outcomes.

Prakash et al [11] explore a novel diagnostic approach that is by using a stacked ensemble learning model for predicting pediatric pneumonia. Recognizing the pressing need for efficient and cost-effective diagnostic tools, the proposed methodology integrates deep features extracted from a fine-tuned Xception model, coupled with a stacked ensemble of classifiers. This approach includes diverse classifiers such as Logistic Regression, K-Nearest Neighbors, Random-Forest, XGB, MLP, Nu-SVC, and SVC in a two-stage stacking model. The advantages of the method encompass robust feature extraction, diversity in classifiers, and impressive diagnostic performance metrics. Notably, the model achieves an accuracy of 98.3

According to Sowmya V et al. [12], the last convolution block in every deep learning model, The final dense layer's DenseNet169-Attn, ResNet101V2-Attn, Resnet152V2-Attn, f ResNet50V2-Attn, and Xception-Attn features are concatenated following dimensionality reduction via kernel PCA. Using KNN, SVC, Logistic Regression, Nu-SVC, and XGB Classifier, the level one classifier is a stairway classifier. The two transformations that are utilized to improve channel interdependencies, and modify filter responses are squeeze and excitation. After being transformed, feature maps are squeezed using global average pooling into a $1 \times 1 \times C$ tensor. The next goal in the module is adaptive scaling weights for the channels. The excitation and squeezing block's excitation module has fully connected layer structures that transfer the scaling weights to outputs based on attention.

Shamrat et al.[13] have extensively utilized transfer learning models like DenseNet121, AlexNet, InceptionV3 MobileNetV2, and VGG19. It focuses on a study introducing a fine- tuned MobileLungNetV2 model for enhanced accuracy in diagnosing various lung conditions. MobileNetV2, known for its efficiency, serves as the base model and is fine-tuned to the specific dataset, achieving an outstanding classification accuracy of 96.97

Nalluri et.al [14] have concluded that an ensemble of classifiers method will be constructed for precise pneumonia detection. We use deep learning methods like DBN, optimizable RNN, and Bi-GRU by modeling them. These classifiers have been trained with an optimal feature that is chosen. Here, optimized RNN, Bi-GRU, and the mean of DBN classifiers are taken. Another point is that , the predicted model of AHGOA was used to fine-tune the RNN's weight to increment the accuracy of detection.

Mahir Kaya et al [15] argue In the first stage, all of the model layers to the final convolution layer which includes the weights acquired from the prior training on the ImageNet dataset—were kept. The weights are used again on a state-of-the-art CNN model to retrain. To improve models for pneumonia identification, further layers are added after the final convolution layer. These extra layers have a Sigmoid activation function for binary classification and have a single-neuron layer, the Dropout layer with a dropout rate of 0.5, and a fully linked dense layer that has 1,024 neurons. In the second stage, These chosen CNN models' fully linked layer features are combined via a method known as feature fusion. From each CNN model's original 1,024 feature vectors, 100, 200, 300, 400, and 500 features were successively chosen using the mRMR and Chi-Square feature selection techniques. Grid Search is used to fine-tune the parameters of Support Vector Machine (gamma and C), RF, and KNN to maximize the performance of the ML methods.

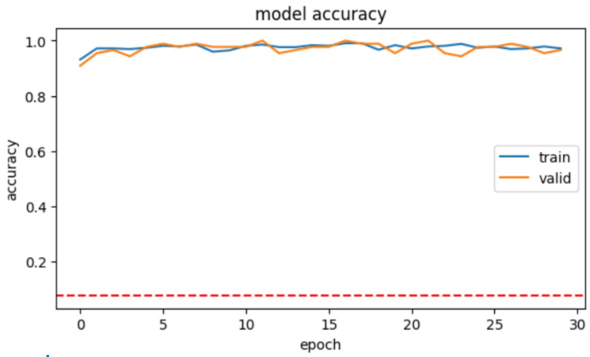
III. Dataset Description

The dataset ,comprising of 5,232 images,is a pediatric radiograph dataset. The chances are either healthy or have a case of either viral or bacterial pneumonia. There are two sections to the dataset: training and testing. 5,232 chest X-rays of kids make up the training set; 3,883 of them are classified as having pneumonia, while 1,349 are not. 2,538 of the photos with the diagnosis of pneumonia are categorized as bacterial, and 1,345 as viral. The test set consists of 234 photos showing no pneumonia and 390 images showing pneumonia, of which 242 are caused by bacteria and 148 by viruses.

IV. Methodology

A. *Finding Out The Best Model*

The methodology employed eight deep learning models, including InceptionV3, ResNet50, MobileNetV2, NASNetMobile, VGG16, Xception, and EfficientNetB0. To train these models, we employed different activation functions such as ReLU, Sigmoid, and Softmax.Training of each model was conducted over a range of epochs, typically varying from 20 to 30 epochs. We evaluated the performance of each model using metrics such as accuracy, precision, recall, and F1 score.



Only loss plots

Fig. 1. Xception Model Loss With Training and Validation Data

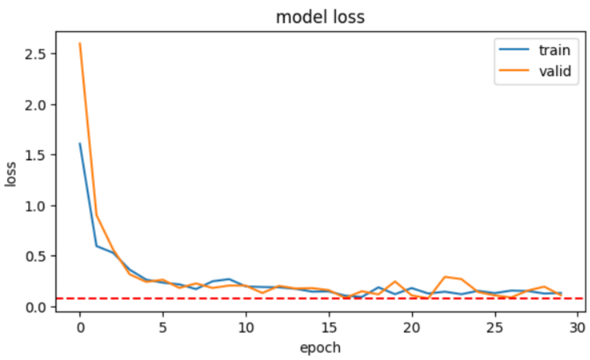


Fig. 2. Xception Model Accuracy With Training and Validation

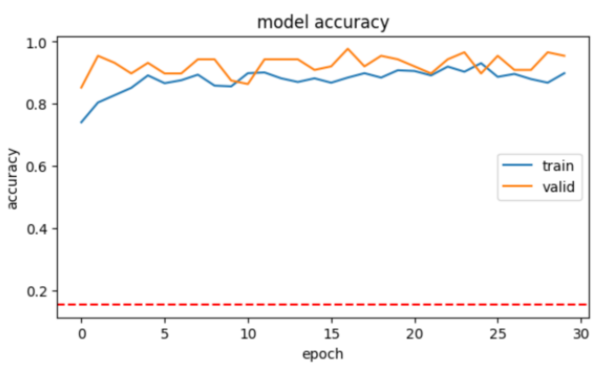


Fig. 3. DenseNet50 Model Loss With Training and Validation Data

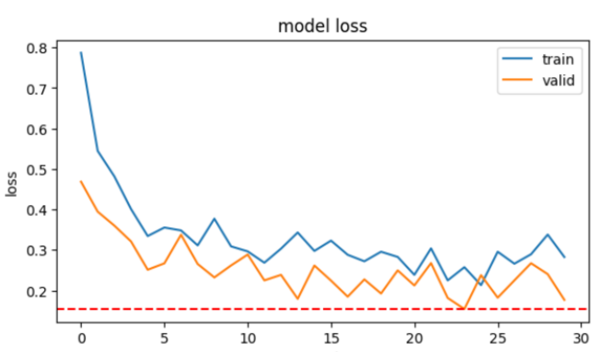


Fig. 4. DenseNet50 Model Accuracy With Training and Validation Data

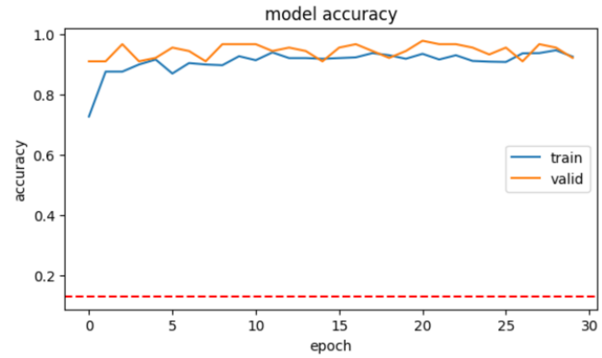


Fig. 5. MobileNetV2 Model Loss With Training and Validation Data

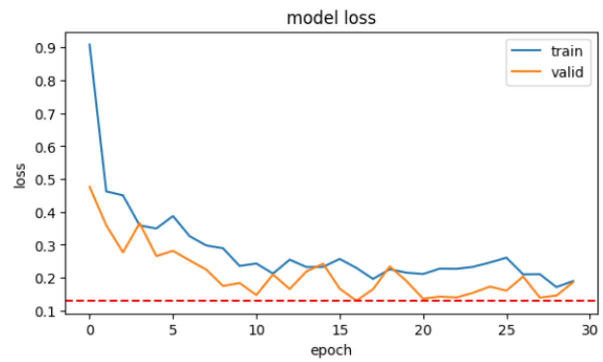


Fig. 6. MobileNetV2 Model Accuracy With Training and Validation Data

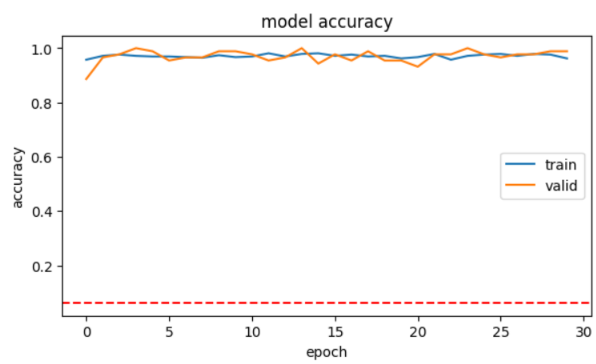


Fig. 7. InceptionV3 Model Loss With Training and Validation Data

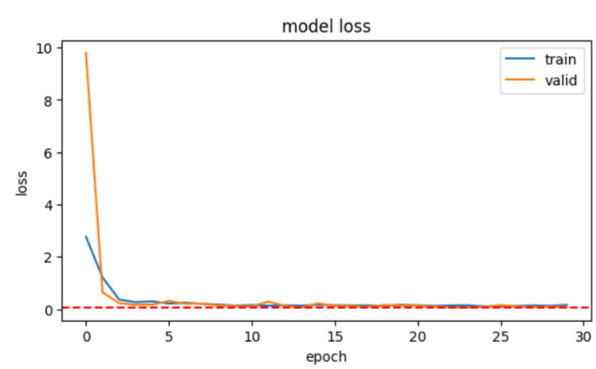


Fig. 8. InceptionV3 Model Accuracy With Training and Validation Data

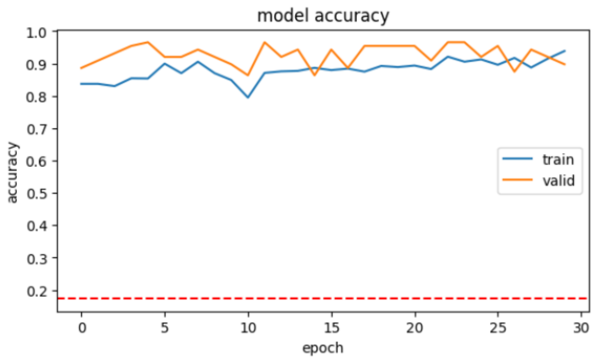


Fig. 9. NasNetMobile Model Loss With Training and Validation Data

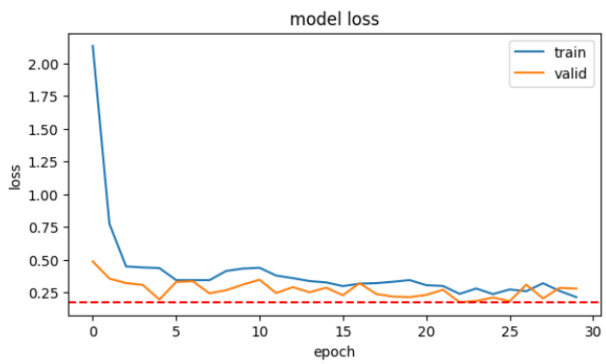


Fig. 10. NasNetMobile Model Accuracy With Training and Validation Data

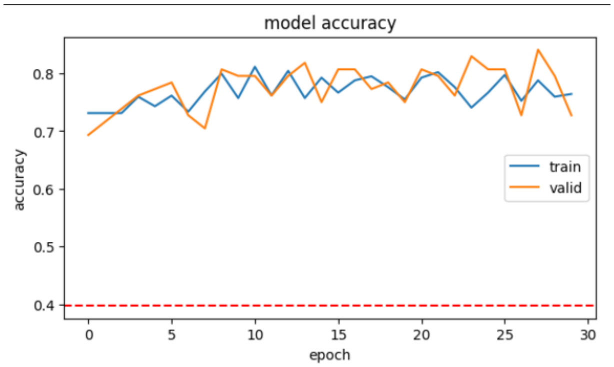


Fig. 11. ResNet50V2 Model Loss With Training and Validation Data

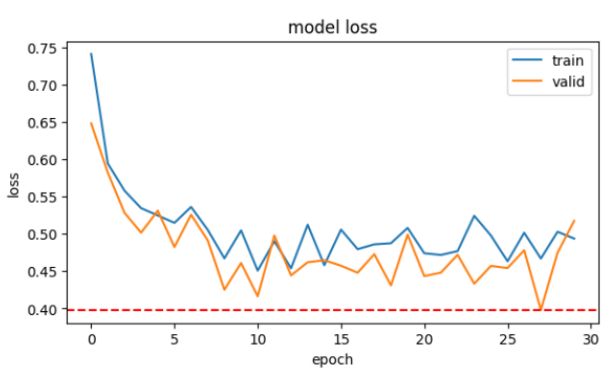


Fig. 12. ResNet50V2 Model Accuracy With Training and Validation Data

TABLE I
ALL MODELS COMPARATIVE ANALYSIS

Model	Accuracy	Precision	Recall	F1-Score	Epochs
Xception	87.66%	83.94	99.23	90.95	30
Densenet50	91.34%	90.77	95.89	93.26	30
MobileNetV2	80.06%	79.81	94.12	91.88	30
InceptionV3	90.87%	87.27	98.46	92.53	30
NasNetMobile	91.02%	93.04	92.56	92.89	30
ResNet50V2	89.74%	86.98	85.81	86.30	30
EfficientNetB0	92.21%	83.91	88.12	78.58	30
VGG16	90.70%	87.55	99.23	93.02	30

B. Training of New Activation Functions

After comparison , We have found that NasnetMobile has the best accuracy and f1 Score. We will train 6 new activation functions like Relu6, Selu , Swish , HardSwish , Gelu and Mish on the Same NasnetMobile Model.

1) ReLU6: It is a direct Modification of the existing "Rectified Linear Unit" Activation Function. Its maximum size to limit the activation can be 6.

$$Y = min(max(0, x), 6) \tag{1}$$

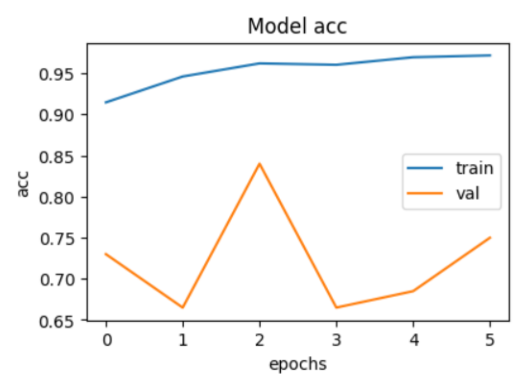


Fig 13. Relu6 - Accuracy

2) SeLU: To get a gradient higher than one for positive inputs, the SELU activation function essentially multiplies scale with the ELU function's output.As long as the weights are properly initialized, the values of scale and alpha are taken to maintain the mean and variance of the inputs between two successive layers of the model.

$$\text{SELU}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leq 0 \end{cases} \tag{2}$$

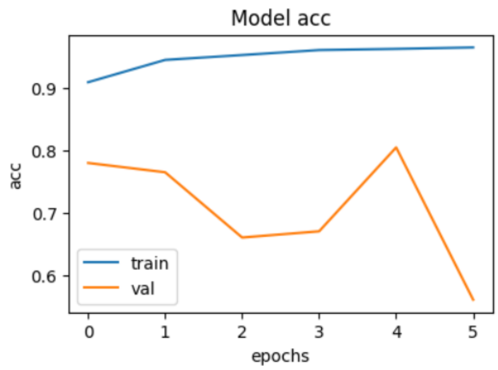


Fig 14. SeLU - Accuracy

3) Mish: It is a smooth, non-monotonic function used as an alternative to conventinal activations like ReLU due to its smoother gradients, leading to better convergence and effiecent performance.

$$f(x) = x \cdot \tanh(\ln(1 + e^x)) \tag{3}$$

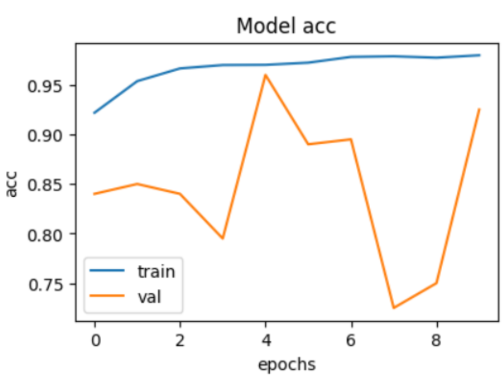


Fig 15. Mish - Accuracy

4) GeLU: It is made of non-linearity weights which input by their percentile, rather than gates input by their sign. It exhibits smoothness and non-linearity, making it suitable for deep learning architectures.

$$f(x) = 0.5x \left(1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right) \tag{4}$$

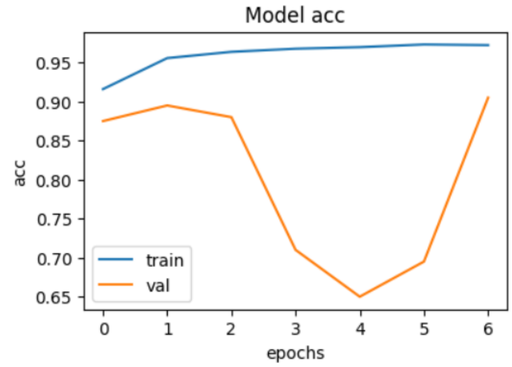


Fig 16. GeLU - Accuracy

5) HardSwish: It replaces the sigmoid component of ReLU Activation Function on a linear scale, simplifying calculations without wasting much in terms of time and performance.

$$f(x) = x \cdot \text{ReLU}\left(\frac{x+3}{6}\right) \tag{5}$$

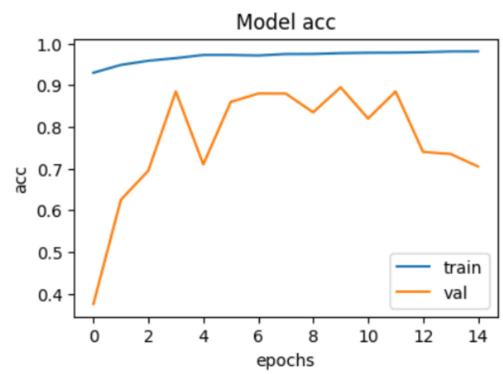


Fig 17. Hardswish - Accuracy

6) Swish: It does not use a Learnable parameter.Its smoothness and non-linearity encourage feature learning, often leading to improved model performance.

$$f(x) = x \cdot \sigma(x) \tag{6}$$

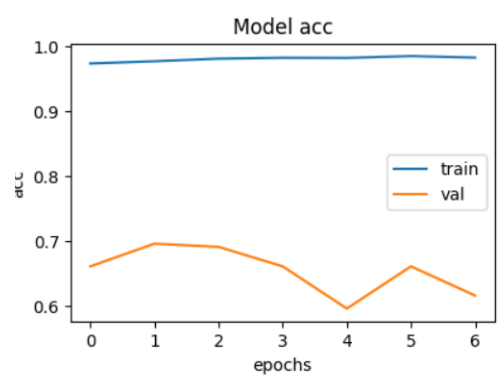


Fig 18. Swish - Accuracy

V. Results

All Activation Functions Comparative Analysis

Activation function	Accuracy	Precision	Recall	F1-Score	Epoch
ReLU6	74.32%	71.11	98.98	82.76	30
SeLU	62.30%	62.30	58.91	76.77	30
Mish	88.27%	84.59	99.23	91.33	30
GeLU	85.30%	81.04	99.74	89.42	30
HardSwish	65.86%	64.67	53.92	78.54	30
Swish	62.82%	62.70	78.44	77.07	30

Among all activation functions , it is GeLU which exhibits better performance . Its accuracy and F1 Score clearly indicates it is better used in Medical Operations for its efficient work.

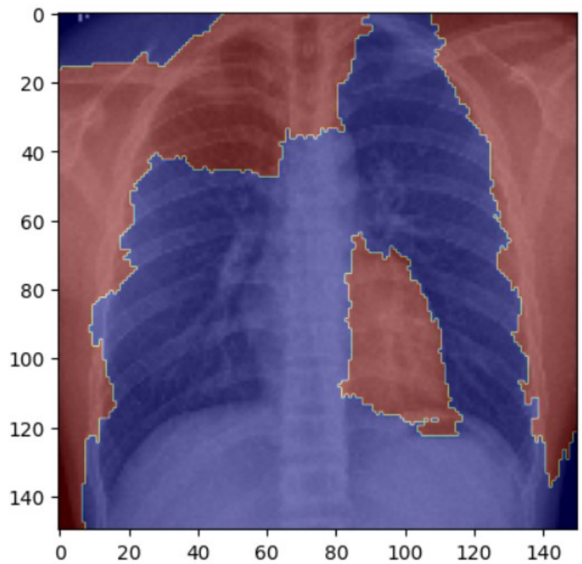


Fig 17. Lime Image using NasNetMobile

We applied lime to see which top 5 segments in the image. We clearly can observe better results in a lime image inspection.

VI. Conclusion

While each activation function exhibits unique characteristics and performance across different tasks, GELU stands out for its combination of smoothness, non-monotonicity, and computational efficiency. GELU's estimation of the Gaussian cumulative distribution function gives desirable qualities, making it particularly strong in various deep learning architectures, especially in the case of HealthCare operations . GELU's balance between simplicity and computational efficiency makes it as a good choice among activation functions in modern applications.

VII. Future Scope

We can implement new activation functions with every passing year for better results. Infact , we can add new Transformers and add Explainability method to achieve higher efficient model training . The future offers a variety of possibilities when it comes to exploring the deep learning applications.

References

[1]Spencer, D. A., Thomas, M. F. (2014). Necrotising pneumonia in children. Paediatric Respiratory Reviews, 15(3), 240-245.

[2] Chen, B., Li, J., Guo, X., Lu, G. (2019). DualCheXNet: dual asymmetric feature learning for thoracic disease classification in chest X-rays. Biomedical Signal Processing and Control, 53, 101554.

[3] Labhane, G., Pansare, R., Maheshwari, S., Tiwari, R., Shukla, A. (2020, February). Detection of pediatric pneumonia from chest X-ray images using CNN and transfer learning. In 2020 3rd international conference on emerging technologies in computer engineering: machine learning and internet of things (ICETCE) (pp. 85-92). IEEE.

[4] Liang, G., Zheng, L. (2020). A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Computer methods and programs in biomedicine, 187, 104964.

[5] Habib, N., Hasan, M. M., Reza, M. M., Rahman, M. M. (2020). Ensemble of CheXNet and VGG-19 feature extractor with random forest classifier for pediatric pneumonia detection. SN Computer Science, 1, 1-9.

[6] Jain, R., Nagrath, P., Kataria, G., Kaushik, V. S., Hemanth, D. J. (2020). Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning. Measurement, 165, 108046.

[7] Guan, Q., Huang, Y. (2020). Multi-label chest X-ray image classification via category-wise residual attention learning. Pattern Recognition Letters, 130, 259-266.

[8] Alhudhaif, A., Polat, K., Karaman, O. (2021). Determination of COVID-19 pneumonia based on generalized convolutional neural network model from chest X-ray images. Expert Systems with Applications, 180, 115141.

[9] Wang, H., Wang, S., Qin, Z., Zhang, Y., Li, R., Xia, Y. (2021). Triple attention learning for classification of 14 thoracic diseases using chest radiography. Medical Image Analysis, 67, 101846.

[10] Fernandes, V., Junior, G. B., de Paiva, A. C., Silva, A. C., Gattass, M. (2021). Bayesian convolutional neural network estimation for pediatric pneumonia detection and diagnosis. Computer Methods and Programs in Biomedicine, 208, 106259.

[11] Prakash, J. A., Ravi, V., Sowmya, V., Soman, K. P. (2023). Stacked ensemble learning based on deep convolutional neural networks for pediatric pneumonia diagnosis using chest X-ray images. Neural Computing and Applications, 35(11), 8259-8279.

[12] Asswin, C. R., KS, D. K., Dora, A., Ravi, V., Sowmya, V., Gopalakrishnan, E. A., Soman, K. P. (2023). Transfer learning approach for pediatric pneumonia diagnosis using channel attention deep CNN architectures. Engineering Applications of Artificial Intelligence, 123, 106416.

[13] Shamrat, F. J. M., Azam, S., Karim, A., Ahmed, K., Bui, F. M., De Boer, F. (2023). High-precision multiclass classifi- cation of lung disease through customized MobileNetV2 from chest X-ray images. Computers in Biology and Medicine, 155, 106646.

[14]Nalluri, S., Sasikala, R. (2024). Pneumonia screening on chest X-rays with optimized ensemble model. Expert Systems with Applications, 242, 122705.

[15] Kaya, M. (2024). Feature fusion-based ensemble CNN learning optimization for automated detection of pediatric pneumonia. Biomedical Signal Processing and Control, 87, 105472.