International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 11| Nov 2023 www.irjet.net p-ISSN: 2395-0072

Deep Learning for Pneumonia Diagnosis: A Comprehensive Analysis of CNN and Transfer Learning on Chest X-rays

Rohith Singhu[1]

[1] *Student, Department of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravathi, Andhra Pradesh, India.*

---***---

Abstract - Pneumonia is a condition characterised by inflammation in the air sacs of one or both lungs, where the affected areas may be filled with either fluid or purulent material, such as pus. Particularly in infants, young children, and people over 60, pneumonia can be fatal.Every year, 2.5 million people die from pneumonia worldwide. A chest x-ray is a crucial test for identifying pneumonia.Chest x-rays can show areas of consolidation that are opaque. However, it takes more time and is less accurate for a professional to diagnose pneumonia using chest X-ray images. The paper suggests the utilisation of convolutional neural network models by medical professionals to accurately detect pneumonic lungs in chest X-rays for the effective diagnosis and treatment of pneumonia. The two primary steps of the suggested framework are the stage of image preprocessing and the stage of feature extraction and image classification. A model that can distinguish between normal and abnormal (pneumonia) images has been developed. As a result, the suggested CNN model-based pneumonia detection method has demonstrated considerable efficiency across all performance metrics, which can help deliver effective patient care and lower death rates.

***Key Words*: Pneumonia, Convolutional Neural Network (CNN), Chest X-ray, Image classification**

1.INTRODUCTION

An infection known as pneumonia causes inflammation in one or both of the lungs' air sacs. The air passages may enlarge with liquid or mucus (purulent material), which can cause a high temperature, chills, a cough that produces pus or phlegm, and difficulty breathing. Pneumonia may be caused by a wide range of organisms, such as pathogens, viruses, and fungal infections. [1] Most occurrences of pneumonia happen in developing and underdeveloped countries because of adverse environmental circumstances, pollution, overcrowding, and a lack of medical facilities. Consequently, early detection and treatment can significantly help prevent the illness from becoming fatal. Radiological inspection of the lungs using computerised tomography (CT), magnetic resonance imaging (MRI), or radiography (X-rays) is commonly used for the identification of lung problems. Non-intrusive and fairly affordable X-ray imaging is used to evaluate the lungs. Demonstrating the conditions of a healthy person as well as the lungs affected by pneumonia is shown in the following figure

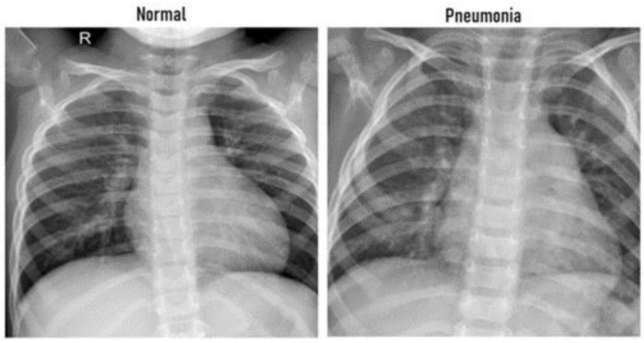

Fig -1: Chest X-ray images without pneumonia and with pneumonia

White dots that might be seen in the chest cavity show that pneumonia has impacted the lungs. [2] Early detection of a deadly disease like pneumonia is always preferable, and if a bacterial infection is likely to be the cause of mild pneumonia, it may be typically treated at home with rest, antibiotics, and lots of fluids. Hospital care can be required in more severe situations. [3]

Deep convolutional neural networks are being used in a variety of current research projects related to medical picture processing. Deep learning employs a variety of models to extract information from pictures presented to the model. Deep learning is currently frequently used in the field of medicine to identify and diagnose diseases and then classify them into a specific disease group. A convolutional neural network is the most extensively used model for medical image processing.

[4] Furthermore, DL models outperformed standard approaches utilising chest X-ray pictures from pneumonia patients. [5]. Deep convolutional neural networks are being used in a variety of current research projects related to medical picture processing. Consequently, the need for an automated pneumonia detection technique arises. The quantity of our data, which allows us to test and evaluate our models in various situations, is a crucial strength of our work in comparison to earlier studies. Transfer learning is a remedy for this hurdle. This approach makes use of the

network parameters generated by a model that was trained on a large dataset to solve a problem requiring a smaller dataset. [6].

CNN's MobileNet, DenseNet, VGG19, InceptionV3, ResNet18, and Efficentnet techniques were trained to identify pneumonia in chest X-ray pictures in order to achieve this aim. We chose to utilise these models to develop the suggested ensemble learning technique, which combines two methodologies that have previously yielded promising results. A solution based on the best ensemble model was created. The dataset in the study, which included chest X-rays from 1,485 instances of viral pneumonia and 1,575 normal cases, was augmented with data mining techniques, and the model developed therein achieved a 97.94% accuracy rate.

2. LITERATURE SURVEY

Medical image classification is seeing an increase in the use of deep learning techniques. This is primarily because these algorithms have a high success rate. Low levels of precision, which were mostly caused by the extraction of feature layer capacity, were one of the difficulties of the initial machine learning models.

In order to attain remarkable test data accuracy or unknown data accuracy, a comparison between previously published pertinent research papers is made in this paper and completed with an ensemble deep learning model.

[7] Arun Sharma, Sheeba Rani, and Dinesh Gupta, "Artificial Intelligence-Based Classification of Chest X-Ray Images into COVID-19 and Other Infectious Diseases.To increase the amount of data and create generic models, they have applied 25 different kinds of augmentations to the original image.

[8] In 2018, Chan YH, Zeng YZ, Wu HC, Wu MC, and Sun HM worked on "Effective pneumothorax detection for chest X-ray images. To identify common pneumothax, SVM is applied, and from lung visuals, features are retrieved using a local binary sequence.

[9] Recently, Mohammed Aledhari, Shelby Joji, Mohamed Hefeida, and Fahad Saeed (2017) proposed an "optimised CNN-based diagnosis system to detect pneumonia from chest radiography using the following three kinds of models: Inception v3, ResNet-50, and VGG16.

[10] Pranav Rajpurkar, Jeremy Irvin, and others recently (2017) looked at this dataset and found that it was better than radiologists at detecting pneumonia. They called their model ChexNet and used the DenseNet-121 layer architecture to find all 14 diseases in the 112,200 images in the dataset. Benjamin Antin et al. (2017) implemented a logistic regression model for pneumonia detection using a similar dataset as the CheXNet mode.

We used CNN's, Mobile Net, Dense Net, VGG19, Inception V3, ResNet18, and Efficentnet techniques. It has been shown that the Efficient Net can range from EfficientNet-B0 to EfficientNet-B7, and that as the model number rises, the model's accuracy rises as well. However, when the number of parameters rises simultaneously by millions.

3. PROPOSED ARCHITECTURE

The following subsections provide detailed coverage of the two key phases of system operations:

Within this segment, we provide a concise overview of the dataset employed in this study and outline the diverse procedures implemented prior to the training of the proposed architectures. Subsequently, we introduce the suggested frameworks designed in order to categorise images generated by X-rays of the chest aimed at pneumonia detection.

3.1 About the Dataset

This study is done using Kaggle Chest X-Ray Images (Pneumonia) [11] for the classification process. The Kaggle Chest X-Ray Images (Pneumonia) dataset comprises 5856 chest X-ray images categorised into two classes: normal and pneumonia

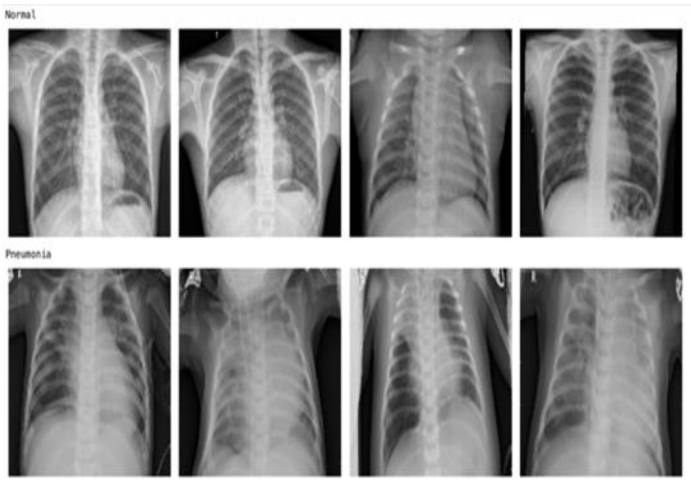

Fig -1: Data samples from the dataset shows normal Cases and shows bacterial pneumonia cases.

3.2 Data pre-processing

In order to standardise images, data pre-processing is being used. In the current instance, this is done by making all chest X-ray images identical in size by 224 x 224 pixels.

Data augmentation is also used to increase the size of the training dataset and improve the model's generalisation ability. Batches of picture information are created using the Keras class image data generator, which applies real-time augmentation methods including random flipping for the training dataset and a range of 0.1 for each dimension shift. This is done to take into consideration the fact that the lungs may not reside at the image's centre.

Fig -3 : shows the image after applying Random Flip.

3.3 Transfer learning

Transfer learning is a technique that lowers learning costs by applying knowledge from a single domain to another. [12]. This can be done by fine-tuning a pre-trained model, where the uppermost layers undergo training and the first layers become frozen. By doing this, the model is able to extract general characteristics from the first several layers.

There are two main steps in the normal process. First, the trained CNN model extracts features from chest X-ray pictures by using relevant patterns and attributes. Frequently, the concluding layers of the pre-trained model, responsible for the original classification task during training, are either eliminated or adjusted. Subsequently, the model undergoes fine-tuning with a dataset specifically crafted for pneumonia detection. This fine-tuning facilitates the adaptation of the model's acquired features to the subtleties inherent in the patterns associated with pneumonia in medical images.

Transfer learning can be effectively employed in pneumonia detection using various CNN architectures such as MobileNet-V2, Inception-v3, ResNet-50, EfficientNet-V1, and EfficientNet-V2. The link between transfer learning and these CNN architectures lies in the utilisation of pre-trained models on large-scale datasets for general image recognition tasks.

Mobilenet-V2:

Across a wide variety of model sizes, MobileNetV2, a cutting-edge mobile model, enhances performance on several tasks and evaluations. A first completely convolutional layer with 32 filters makes up MobileNetV2's architecture. It then follows with 19 residual bottleneck layers. [13] The ReLU6 non-linearity is employed due to its resilience in low-precision computations.The network uses a kernel size of 3 x 3, which is standard for modern networks, and also utilises dropout and batch normalization during training. The network uses a constant expansion rate, except for the first layer. Experiments revealed that performance is equivalent for growth rates around 5 and 10, with bigger networks performing better at greater expansion rates and smaller ones performing better at lower expansion rates. Transfer learning involves taking a pre-trained MobileNet-V2 model, which has been trained on diverse image datasets, and adapting it to the pneumonia detection task. The pre-trained MobileNet-V2 serves as a feature extractor, capturing relevant patterns from chest X-ray images.

Inception-v3:

The Inception architecture, as used in GoogLeNet [14], was designed to perform well under memory and computational constraints. In comparison to VGGNet, the Inception network is more economical because of both the total number of parameters used and the computing productivity. However, the methods used to achieve this efficiency add additional complexity to the architecture, making it harder to make changes to the network and adapt it to new use cases while maintaining its efficiency [15]. To address this, the Inception v3 It makes use of methods including regularization, reduction of dimensions, simplified convolutions, and parallelized computations. to make it easier to adapt the network. Similarly, Inception-v3 can be employed in transfer learning for pneumonia detection. Convolutional layers of the previously trained Inception-v3 model are employed as a foundation to extract relevant characteristics from X-ray pictures.

Resnet-50:

ResNet-50 is a sophisticated deep convolutional neural network architecture that incorporates residual connections, allowing the network to skip one or more layers during the training process. In the written work, it was first mentioned,Deep Residual Learning for Image Recognition" (He et al., 2015). [16] and has 50 weight layers. The network can learn deep architectures thanks to the leftover connections without running into the issue of vanishing gradients. The ResNet-50 architecture is constructed by stacking multiple residual blocks, each with several convolutional layers. The ResNet-50 architecture has been widely adopted in computer vision tasks such as object detection, semantic segmentation, and many more. The residual connections and deep architecture make ResNet-50 a powerful model for image classification tasks, and it can also be used as a feature extractor for other tasks. Transfer learning with ResNet-50 involves leveraging the knowledge gained from a pre-trained model on a broad set of images. The deep residual networks of ResNet-50 can capture intricate features in chest X-ray images, enhancing the model's ability to detect pneumonia.

Efficientnet, V1

Google Research created the EfficientNet family of image categorization models with the intention of increasing convolutional neural networks' (CNNs') effectiveness by scaling up the network's architecture while also reducing the number of parameters and computational costs. EfficientNet-V1 is the first version of this family, and it was introduced in a paper by Mingxing Tan and Quoc V. Le, published in 2019. [6] The authors showed that EfficientNet-V1 outperforms other cutting-edge models on a variety of image categorization benchmarks, including ImageNet and COCO, while also being considerably better in regards to the quantity of parameters and computing expenses. Both versions of EfficientNet can be utilised in transfer learning for pneumonia detection. These models, known for their efficiency and effectiveness, provide a strong foundation for feature extraction, and fine-tuning them on a pneumonia-specific dataset refines their capabilities for the targeted medical image classification.

EfficientNet-V2

EfficientNet-V2 is a version of the EfficientNet architecture that was introduced by Mingxed Tan, Ruoming Pang, Vijay Vasudevan, and Quoc V. Le in 2020.

In this research, we proposed several modifications to the original EfficientNet-V1 architecture in order to further improve its efficiency and performance. The main modification is the use of a new scaling method called "linear scaling rule," which enables scaling up the network's width and resolution by a fixed factor while also adjusting the depth proportionally. [17] This results in a more efficient use of computation and a better balance between the width, depth, and resolution of the network.

4. PROPOSED MODEL ARCHITECTURE

Our presented model involves the use of the convolutional neural network (CNN) architecture. In the provided neural network architecture, the model begins with a Conv2D layer with 8 filters of size 3x3, resulting in an output shape of (222, 222, 8). Subsequently, another Conv2D layer is applied, maintaining the same filter size and resulting in an output shape of (220, 220, 8). A MaxPooling2D layer having a pool dimension of 2x2 is used after every convolutional layer to decrease the spatial dimensions. The process is repeated with two more sets of Conv2D and MaxPooling2D layers, resulting in a final output shape of (53, 53, 8).

The model then undergoes a flattening operation to transform the 3D output into a 1D array of size 22472. Dropout regularisation is implemented after flattening. During training, a random percentage of input units is set to zero to help minimise overfitting. Finally, a dense layer with two output units and a sigmoid activation function is utilised for binary classification between the normal and pneumonia classes.

The model architecture comprises a total of 46,922 parameters, all of which are trainable. For effective gradient-based optimization during training, the selected Adam optimizer is used, and the function of binary cross-entropy loss is utilized, which is appropriate for binary classification tasks. This architecture demonstrates a sequential arrangement of convolutional and pooling layers, followed by flattening and fully connected layers, providing a comprehensive structure for image classification tasks.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 8)	224
conv2d_1 (Conv2D)	(None, 220, 220, 8)	584
max_pooling2d (MaxPooling2D)	(None, 110, 110, 8)	0
conv2d_2 (Conv2D)	(None, 108, 108, 8)	584
conv2d_3 (Conv2D)	(None, 106, 106, 8)	584
max_pooling2d_1 (MaxPooling2	(None, 53, 53, 8)	0
flatten (Flatten)	(None, 22472)	0
dropout (Dropout)	(None, 22472)	0
dense (Dense)	(None, 2)	44946

Total params: 46,922
Trainable params: 46,922
Non-trainable params: 0

Fig -4: Summary of CNN Model

The proposed CNN model is trained on a dataset specifically curated for pneumonia detection, ensuring a balanced representation of both classes. During training, the model was trained to distinguish between chest X-ray pictures without pneumonia and normal.

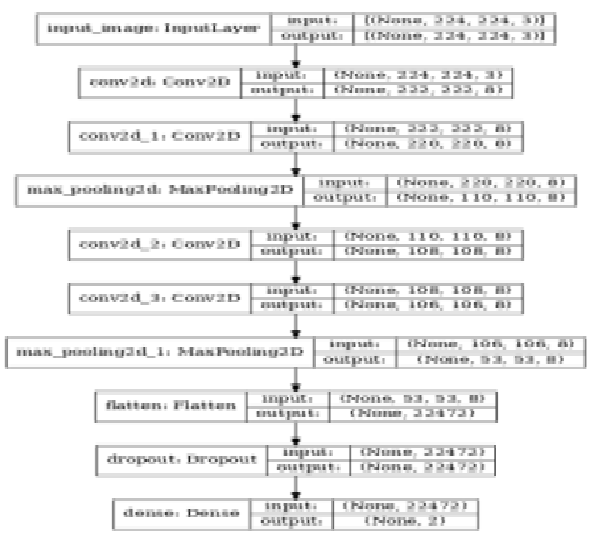

Fig -5 : Proposed Model

5. RESULTS AND DISCUSSION

As stated in (1), (2), (3), (4), and (5), the models are assessed using the performance metrics accuracy, recall, precision, F1 score, and Mathews coefficient. True positive, true negative, false positive, and false negative are represented by the symbols TP, TN, FP, and FN, correspondingly.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FN+TN+FP)} \quad (1)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (2)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (3)$$

$$\text{F1score} = 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} \quad (4)$$

Mathew Correlation Coefficient =

$$\frac{(TP \times TN) - (FP \times FN)}{(TP+FP)(TP+FN)(TN+FP)(TN+FN)} \quad (5)$$

	accuracy_score	precision_score	recall_score	f1_score	matthews_corrcoef
model_cnn	0.95849	0.958596	0.958491	0.958539	0.89249
model_mv2	0.95283	0.952570	0.952830	0.952661	0.87676
model_resnet	0.95660	0.956604	0.956604	0.956604	0.88733
model_efficientnet_v1	0.95472	0.955632	0.954717	0.955024	0.88444
model_efficientnet_v2	0.94717	0.950014	0.947170	0.947913	0.86856

Fig -6: documents the performance metrics recall, precision, F1 score, and Matthews correlation coefficient for the five models with the highest accuracy

The proposed CNN model is trained over a 30-epoch period using an average batch size of 32 in the training process, 131 steps per epoch at training, then 32 for the testing phase, and 20 steps per epoch in testing.The graph shows the training accuracy and validation obtained at each epoch throughout the training process. We can observe that training stops at 25 epochs; this is due to the utilisation of early stopping regularization. The graph above demonstrates how the validation accuracy and training accuracy are nearly identical. Consequently, our deep learning model appears to be free from overfitting. Training for 25 of these epochs therefore produced optimal outcomes. with a high accuracy of 95.54% and a minimal loss of 0.1237 for training and a corresponding 95.0% accuracy and 0.1419 loss for validation.

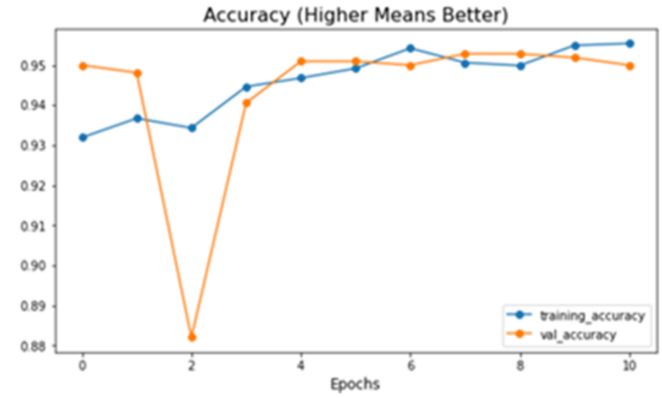

Chart -1: Training and validation accuracy

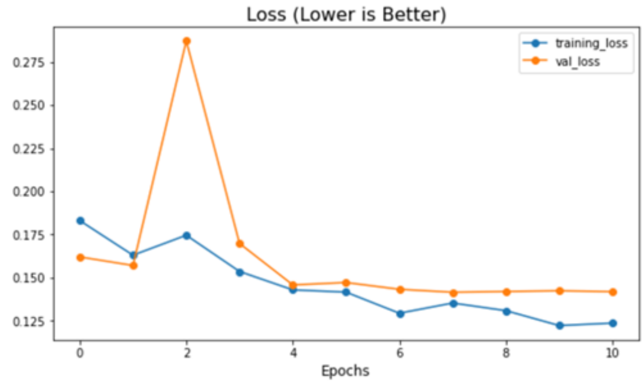

Chart -2: Training and validation loss

The graphical representation reveals that MobileNetV2, EfficientNetV1, and our devised CNN model exhibit superior accuracy. MobileNetV2 attains an accuracy of approximately 95% with remarkable efficiency, requiring a mere 15 epochs for convergence. In contrast, both EfficientNetV1 and our CNN model demonstrate comparable accuracy, achieving around 95% accuracy, yet necessitating a more extended training period of 25 epochs to reach this performance threshold.

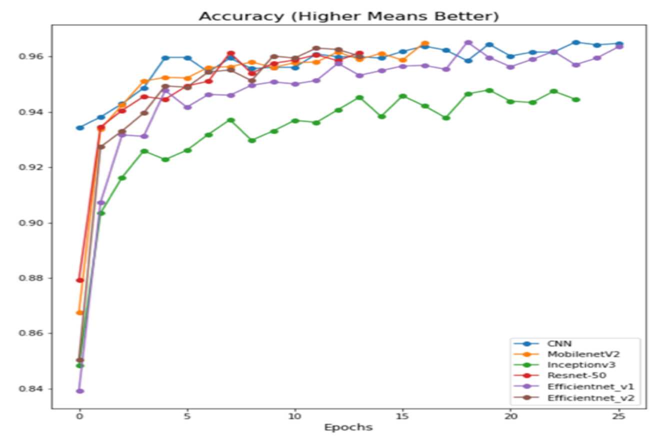

Chart -3: Variation of accuracy rates achieved by CNN, MobilenetV2, Inceptionv3, Resnet-50, EfficientnetV1 and EfficientnetV2

In our context, sensitivity refers to the percentage of individuals testing positive for pneumonia, while specificity pertains to the percentage of individuals without pneumonia who receive negative test results.

Sensitivity = TP / (TP + FN) - (5)

Specificity = TN / (FP + TN) – (6)

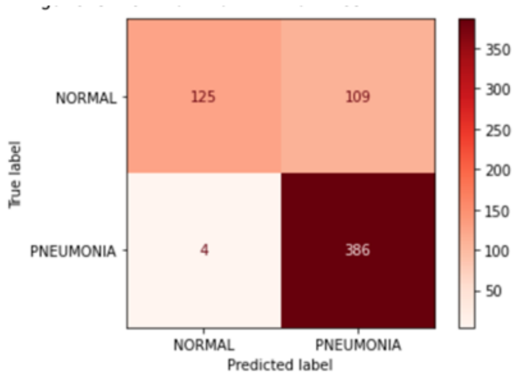

Fig -8: Shows the CNN model's confusion matrix

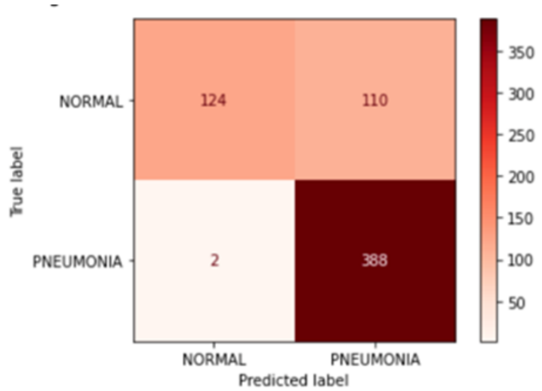

Fig -9: illustrates MobileNet's confusion matrix

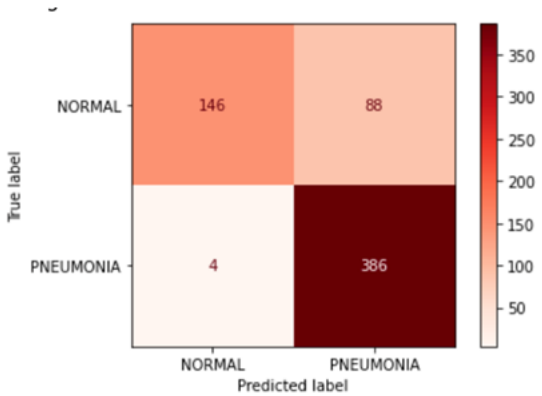

Fig -10: shows the Efficientnet confusion matrix

Hence, we calculate from Eqs. (5) and (6) that the sensitivity for CNN, MobileNetV2, and EfficientNetV1 is 0.9689, 0.9763, and 0.97. The specificities for CNN, MobileNetV2, and EfficientNetV1 are 0.77, 0.77, and 0.81, respectively. Both of these scores are above 75%, indicating that the model works admirably in both areas, properly identifying the majority of positive pneumonia cases and ruling out the negative cases.

6. CONCLUSION

In this paper, our aim is to identify a more straightforward approach for pneumonia detection using chest X-rays (CXRs). We accomplish this by assessing the effectiveness of five different convolutional neural network (CNN) architectures—Mobilenet-V2, Inception-V3, Resnet-50, EfficientNet-V1, and EfficientNet-V2—all trained on the same dataset. Following our evaluation, we choose the most appropriate model based on criteria such as ease of training (lower computational requirements and faster processing), interpretability, and strong performance metrics. The selected model is MobileNetV2, achieving a 95% accuracy level in just 15 epochs. In comparison, EfficientNet-V1 and the traditional CNN model exhibit a comparable accuracy of approximately 95% after 25 epochs, presenting a promising alternative. Future work on detection of pneumonia can be done by developing an intelligent triage system using the Res-CovNet framework. Recent findings from thorough investigations suggest that the Res-CovNet model, designed for diagnosing patients with pneumonia and COVID-19, exhibits a heightened capability to diagnose individuals more rapidly. Importantly, the proposed Res-CovNet framework is highly scalable, featuring broad flexible capacities, making it adaptable for various computer vision applications. Additionally, this framework has the potential to extend its utility to the diagnosis of other life-threatening diseases associated with cancer. Subsequent efforts in the development of this framework will be concentrated in this particular direction.

REFERENCES

[1] A. Ravikumar and H. Sriraman, “Attenuate Class Imbalance Problem for Pneumonia Diagnosis Using Ensemble Parallel Stacked Pre-Trained Models,” Computers, Materials & Continua, vol. 75, no. 1, pp. 891–909, 2023, doi: https://doi.org/10.32604/cmc.2023.035848.

[2] Kallander, D. H. Burgess, and S. A. Qazi, “Early identification and treatment of pneumonia: a call to action,” The Lancet Global Health, vol. 4, no. 1, pp. e12–e13, Jan. 2016, doi: https://doi.org/10.1016/s2214-109x(15)00272-7.

[3] R. S. Fhon, L. M. Silva, Z. E. Leitón-Espinoza, F. de B. Matiello, J. S. de Araujo, and R. A. P. Rodrigues, “Hospital care for elderly COVID-19 patients,” Revista Latino-Americana de Enfermagem, vol. 28, doi: https://doi.org/10.1590/1518-8345.4649.3396.

[4] D. Meedeniya, H. Kumarasinghe, S. Kolonne, C. Fernando, I. D. la T. Díez, and G. Marques, “Chest X-ray analysis

empowered with deep learning: A systematic review," Applied Soft Computing, vol. 126, p. 109319, Sep. 2022, doi: https://doi.org/10.1016/j.asoc.2022.109319.

[5] R. Kundu, R. Das, Z. W. Geem, G.-T. Han, and R. Sarkar, "Pneumonia detection in chest X-ray images using an ensemble of deep learning models," PLoS ONE, vol. 16, no. 9, p. e0256630, Sep. 2021, doi: https://doi.org/10.1371/journal.pone.0256630.

[6] P. Marcelino, "Transfer learning from pre-trained models," Medium, Oct. 23, 2018. https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751

[7] A. Sharma, S. Rani, and D. Gupta, "Artificial Intelligence-Based Classification of Chest X-Ray Images into COVID-19 and Other Infectious Diseases," International Journal of Biomedical Imaging, vol. 2020, pp. 1–10, Oct. 2020, doi: https://doi.org/10.1155/2020/8889023.

[8] Y.-H. Chan, Y.-Z. Zeng, H.-C. Wu, M.-C. Wu, and H.-M. Sun, "Effective Pneumothorax Detection for Chest X-Ray Images Using Local Binary Pattern and Support Vector Machine," Journal of Healthcare Engineering, vol. 2018, pp. 1–11, 2018, doi: https://doi.org/10.1155/2018/2908517.

[9] M. Aledhari, S. Joji, M. Hefeida, and F. Saeed, "Optimized CNN-based Diagnosis System to Detect the Pneumonia from Chest Radiographs," Nov. 2019, doi: https://doi.org/10.1109/bibm47256.2019.8983114.

[10] P. Rajpurkar et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," arXiv.org, 2017. https://arxiv.org/abs/1711.05225

[11] P. MOONEY, "Chest X-Ray Images (Pneumonia)," www.kaggle.com, 2018. https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

[12] N. Donges, "What is transfer learning? Exploring the popular deep learning approach," Built In, Aug. 25, 2022. https://builtin.com/data-science/transfer-learning

[13] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Jun. 2018, doi: https://doi.org/10.1109/cvpr.2018.00474.

[14] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, "Rethinking the Inception Architecture for Computer Vision." Available: https://arxiv.org/pdf/1512.00567v3.pdf

[15] "Inception V3 Model Architecture," OpenGenus IQ: Computing Expertise & Legacy, Sep. 05, 2021. https://iq.opengenus.org/inception-v3-model-architecture/

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," arXiv.org, Dec. 10, 2015. https://arxiv.org/abs/1512.03385

[17] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," arXiv.org, 2019. https://arxiv.org/abs/1905.11946

[18] M. Tan et al., "MnasNet: Platform-Aware Neural Architecture Search for Mobile," arXiv:1807.11626 [cs], May 2019, Available: https://arxiv.org/abs/1807.11626