Comparative Analysis and Validation for Diagnosis of Pneumonia through Chest X-rays using Deep Learning Models

[1]Shivani Pandey, [2]Yashashree Shinde, [3]Shatabdi Pingale, [4]Sakshi Rathi, [5]Sonali Surpatne

[1]Assistant Professor, [2,3,4,5]Student
Computer Engineering,
Pimpri Chinchwad College of Engineering, India.

***Abstract*—Pneumonia has been the widespread disease caused by a respiratory infection and has rapid spread and relatively high mortality rate. It is mostly found in the children of age group of five. Early detection and treatment of pneumonia will significantly reduce its mortality rate. X-ray diagnosis is currently recognized as a relatively effective method to diagnose the pneumonia. An experienced doctor's visual analysis of a patient's X-ray chest radiograph takes about 5 to 15 minutes. When cases are concentrated, the doctor's clinical diagnosis is undoubtedly put under tremendous strain. As a result, relying on the imaging doctor's naked eye has a very low efficiency. So, it is important to use AI to help doctors diagnose pneumonia from clinical images. Furthermore, artificial intelligence recognition is extremely fast, and convolutional neural networks (CNNs) have outperformed humans in image identification. To achieve a solution to the problem we used the Kaggle dataset with chest X-ray images for classification, which included 5216 train and 624 test images and two classes: normal and pneumonia. We conducted studies in which we used five mainstream network algorithms to classify these diseases in the dataset and compared the results, in which custom CNN model achieved a higher accuracy rate than other methods. The accuracy gained was about 92.7 % Additionally, the improved Custom CNN network may be extended to other areas for application and better results.**

***Index Terms*—Pneumonia, Custom CNN, Resent-50, VGG-16, Dense Net, Inception**

I. INTRODUCTION

Pneumonia is inflammation of a tissues in or both of the lungs caused by bacteria, viruses, or fungi in air sacs as well as empyema, a condition in which the lung fills with fluid. It is responsible for more than 15% of all deaths in children under the age of five and also in adults. Pneumonia is most common in developing and underdeveloped countries, where overcrowding, pollution, and unsanitary environmental conditions exacerbate the situation and medical resources are meagre. As a consequence, early diagnosis and management can be considered vital in preventing the disease from becoming fatal. Worldwide, 450 million people get infected with pneumonia, and 4 million dies as a result of the disease. In the United States of America, 1 million people require hospital care each year, and 50 thousand people die from the disease [11]. The numerical difference between infection and death rates demonstrates the importance of early detection of disease. For diagnosis purpose the radiological examinations like computed tomography (CT), magnetic resonance imaging (MRI), or radiography (X-rays) of the lungs are commonly used. X-ray imaging is a non-invasive and relatively low-cost examination of the lungs. Figure 1 shows the pneumonic and a healthy lung X-ray [1]. Infiltrates, which resemble white spots on a pneumonic X-ray and are marked with red arrows, indicate the condition is not healthy. However, chest X-ray examinations for the detection of pneumonia are subordinate to subjective variability. Thus, it is the responsibility of the radiologists to correctly identify the white spots corresponding to pneumonic fluid. The error margin of human eye may fail to make correct diagnosis, therefore an automated system for detecting pneumonia is required. In this study, we developed a computer-aided diagnosis (CAD) system that accurately classifies chest X-ray images using a various deep learning model. Deep learning is a powerful artificial intelligence tool that can help solve a wide range of complex computer vision problems. Deep learning models, specifically convolutional neural networks (CNNs), are widely used for image classification. However, such models function efficiently only when given a large amount of data. Such a large amount of labelled data is difficult to obtain for biomedical image classification problems because it requires expert doctors to classify each image, which is a costly and time-consuming task.

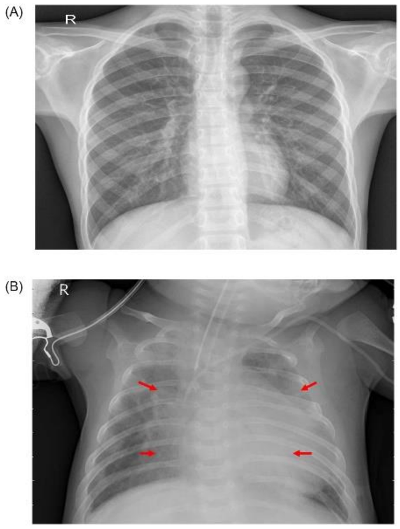

Figure 1 The red marks showing the affected area of the Pneumonia Disease

II.RESEARCH OBJECTIVE

1. As previously stated, pneumonia affects a large number of people, particularly children, primarily in developing and underdeveloped countries characterized by risk factors such as overcrowding, poor sanitary conditions, and malnutrition, as well as a lack of good medical facilities
2. Early detection of pneumonia is critical for complete cure. The most common method of diagnosis is X-ray scan examination, but this is dependent on the radiologist's interpretative ability and is frequently not agreed upon by the Radiologists.

III.ORGANISATION OF PAPER

The section I comprises the paper's introduction, which states the symptoms of pneumonic lungs and how most of the infants get easily affected by it the section II deals with the research objective followed by section IV a literature review, which does include a deep study of various deep learning and machine learning research articles available related to pneumonia detection. The section V incorporates the data and the methods used to achieve the results. Finally, in section VI, the experiments conducted are presented, also the results showing the accuracy, recall and precision of various CNN architecture is displayed in section VII and section VIII concludes the paper by highlighting future scopes.

IV.LITERATURE SURVEY

One challenge in early pneumonia diagnosis is the reliance on human detection. Trained radiologists must have a keen eye to differentiate between the various colors present in the heterogeneous distribution of air in the lungs, which can appear differently on X-ray images but may not indicate pneumonia fluid. This is why it's crucial for radiologists to determine if the white spots on an X-ray film correspond to fluid [11]. Due to the margin of error in human diagnosis, radiologists often fail to correctly identify pneumonia, leading to either false positives or false negatives, both of which have significant consequences for the patient. Computational methods in diagnosis offer a more consistent and reliable alternative. Figure 1 shows various images with and without pneumonia, illustrating the difficulty in distinguishing between healthy and pneumonia images. Some studies have investigated the use of machine learning and heat maps to detect pneumonia through chest x-rays which are graphical representations of temperature or radiation variation over a certain area or period. Computerized lung sound analysis has also been used to differentiate between normal lung function and pulmonary pathology [2], which involves diagnosing neoplastic and non-neoplastic lung diseases. Another method for diagnosing pneumonia caused by fungi, specifically p. carinii pneumonia, involves examining induced sputum and using indirect immunofluorescence [13]. Apart from traditional X-ray imaging, other diagnostic techniques like bronchoalveolar lavage, lung biopsy, and lung ultrasonography are also used for lower respiratory tract infection diagnosis. Neonatal pneumonia, which is characterized by lung consolidation with irregular margins and air bronchograms, pleural line abnormalities, and interstitial syndrome, can also be detected using these methods. Apart from this the published papers helps in better understanding of the topic which includes “Automated Detection of Pneumonia from Chest X-Rays using Convolutional Neural Networks" by Wang et al. (2017) [2]. The authors used a CNN-based approach to detect pneumonia from chest x-rays. They achieved an AUC of 0.92 on a test set of 100 x-rays. The paper “Pneumonia Detection using Convolutional Neural Networks from Chest X-Ray Images" by Rajpurkar et al. (2017) [14]. The authors used a CNN-based approach to classify chest x-rays as normal, showing evidence of pneumonia, or showing evidence of other lung diseases. They achieved an AUC of 0.92 on a test set of 420 x-rays. “Pneumonia Detection using Faster R-CNN with Pyramid Features from Chest X-Ray Images" by Wang et al. (2018) [4]. The authors used a faster R-CNN object detection model with pyramid features to detect pneumonia from chest x-rays. They achieved an AUC of 0.97 on a test set of 585 x-rays. “Pneumonia Detection from Chest X-Ray Images using Deep Learning and Support Vector Machine" by Moradi et al. (2018) [3]. The authors developed a deep learning model based on a CNN and support vector machine (SVM) to diagnose pneumonia from chest x-rays. They achieved an accuracy of 93.5% on a test set of 585 x-rays. “Pneumonia Diagnosis from Chest X-Ray Images using Deep Learning and Fuzzy Logic" by Kumar et al. (2020) [9]. The authors developed a deep learning model based on a CNN and fuzzy logic to diagnose pneumonia from chest x-rays. They achieved an accuracy of 92.5% on a test set of 150 x-rays." Deep Learning for Pneumonia Diagnosis from Chest X-Rays" by Liu et al. (2018) [12]. The authors proposed a deep residual network (ResNet) architecture for pneumonia detection from chest x-rays. They achieved an accuracy of 89.5% on a test set of 5,856 x-rays. There have also been studies that focused on developing deep learning models for other medical image analysis tasks, such as detecting breast cancer using mammography images, detecting skin cancer using mammographic images, and diagnosing diabetic retinopathy using fundus images [8]. These studies have demonstrated the potential of deep learning in medical image analysis and its ability to achieve high accuracy in diagnosis. Furthermore, these models have the potential to assist doctors in making more accurate and efficient diagnoses, ultimately improving patient outcomes.

V.MATERIALS AND METHODS

Data

The dataset was publicly available from gaggle.com. The dataset was a part of Anterior-posterior chest X-ray images chosen from past cohorts of pediatric patients aged between one to five years old at Guangzhou Women and Children's Medical Center in Guangzhou. The chest X-ray scans were conducted during the patients' regular clinical care. This enables automated methods to detect and classify human diseases from medical images. The images in the dataset are varying resolutions such as 1128 × 624 to 1438 ×1260. There are 5863 JPEG X-ray images of two types (pneumonia and normal). Before being analyzed, all chest radiographs were reviewed for quality control, with all low-quality or unreadable images being eliminated. Before training the AI system, the diagnoses for the images were graded by two expert physicians. A third expert also reviewed the evaluation set to account for any grading errors. Firstly, it is important to conduct data analysis and preprocessing. This involves converting the images obtained from the dataset into a NumPy array and resizing them to 226 x 226 and 64 x 64 according to the pre-trained models to increase the amount of data available for training [3][6].

Category	**Test Samples**	**Train Samples**	**Validation Samples**	**File Type**
Normal	1341	234	8	JPEG
Pneumonia	3875	390	8	JPEG

Table 1 Characteristics of Dataset

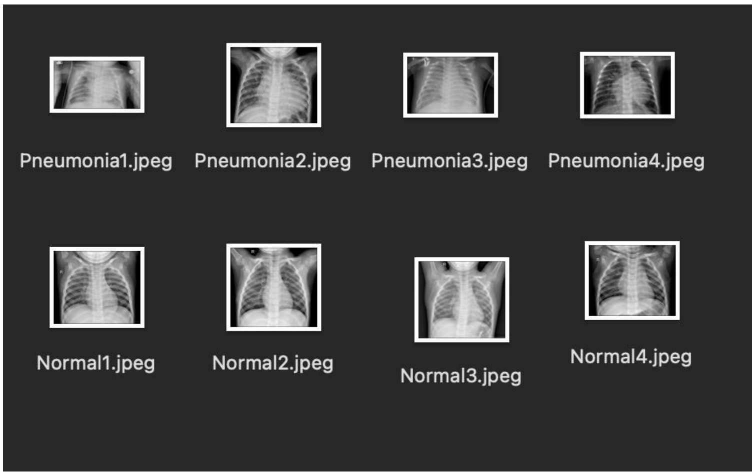

Figure 2 Data Samples form Dataset showing Normal and Pneumonia Patients

Methods

Convolutional Neural Network are the powerful tools for recognizing local patterns in data samples. There are considered as suitable architectures for the classification task. This makes them ideal for applications requiring critical object detection, such as computer vision (CV) activities. The structure of a CNN resembles to the neural structure of the human brain. Similar to how the brain has billions of neurons, CNNs also have neurons, but they are structured differently. In actuality, a CNN's neurons are set up similarly to the frontal lobe of the brain, which processes visual stimuli. This configuration guarantees that the full visual field is covered, avoiding the issue with typical neural networks' partial image processing that requires images to be given to them in low-resolution portions. Here the affected area is treated as an object to be identified. Therefore, to serve the purpose of detecting pneumonic lungs we have used following architecture in this study which are as: InceptionV3, Dense Net, RESNET50, and VGG16 and the custom 5-layer CNN model for the comparison purpose.

1)Inception V3 - InceptionV3, is an instance of deep convolutional neural networks (ConvNets) with a unique architecture that makes it highly efficient for large scale image recognition tasks. It is a complex network with multiple branches and parallel processing paths, each designed to recognize different scales, orientations, and abstraction levels of features in an image. This enables InceptionV3 to perform well even in complex and large-scale image recognition tasks.

2) ResNet 50 - ResNet-50 is a 50-layer convolutional neural network (48 convolutional layers, one MaxPool layer, and one average pool layer). The ResNet-50 architecture is specifically designed to increase the depth of the neural network while reducing the risk of vanishing gradients and overfitting. The residual blocks in the network help to ensure that the output of the previous layer is added to the current layers output, which allows the network to better retain information from earlier in the network. This results in the ResNet-50 architecture outperforming traditional deep neural networks, making it a popular choice for computer vision and image classification tasks.

3) VGG 19 - VGG-19 is part of the VGG network architecture family, which is known for its simple architecture and strong performance in image classification tasks. The architecture consists of several convolutional and max pooling layers, followed by several fully connected layers, and finally, a SoftMax activation function that outputs the class probabilities. The convolutional layers use small filters (e.g.,3x3), which are stacked multiple times to extract hierarchical features from the input image.

4) DenseNet - The DenseNet architectures proposed by Huang et al. provide a rich feature representation while being computationally efficient. The primary reason for this is that, as the feature maps in the current layer are concatenated with those from all preceding layers in each layer of the DenseNet model. Because the convolutional layers accommodate fewer channels, the number of trainable parameters is reduced, and the model is thus computationally efficient. Furthermore, combining the feature maps from previous layers with the current layer improves feature representation.

5) Custom CNN - The custom CNN uses for this proposed study comprises of 5 layers the first one is the input layer; this layer takes the input image and applies a convolutional filter to extract features. The number of filters can be chosen based on the complexity of the input image. It also includes pooling layer and convolutional layer. The last layer is the fully connected layer The output of the previous convolutional layer is flattened and fed into a fully connected layer, which performs the final classification task. The activation function used is ReLU.

VI.EXPERIMENTS

In this section, training strategies and test results were presented. Before the training phase the resizing of the images is done. Also, all the image pixels are normalized range of [0,1]. We have trained all the networks by using the same parameters. Furthermore, in order to facilitate comparison, we set the number uniformly to 50 epochs categorical cross entropy as loss function. The proposed networks are implemented by using keras deep learning framework using Python programming language on a Windows operating system and Google collab repository.

Evaluation Metrics

The accuracy rate provides an overall measure of the model's correct predictions. However, if the dataset is imbalanced, a model's high accuracy rate does not guarantee its ability to distinguish different classes equally. A model that can be generalised to all classes is especially important in medical image classification. In such cases, the "precision" and "recall" values provide information about the model's performance. "Precision" displays the model's positive label prediction accuracy. This is the ratio of correct predictions to total predictions produced by the model. In contrast, "recall" measures the percentage of true positives predicted by the model. These two metrics are used to determine whether the model can reduce the number of FP and FN predictions. "F1-Score" strikes a balance between "precision" and "recall," considering both False Positive's and False Negative's. It penalizes extreme values of "precision" and "recall," which are obtained at the expense of the other. Thus, in medical image classification, it is beneficial to consider evaluation metrics rather than just the accuracy rate in order to obtain a precise identification of both a healthy and a diseased person.

$$\text{Accuracy} = (TP+TN)/(TP+FN+TN+FP) \quad (1)$$
$$\text{Precision} = TP/(TP+FP) \quad (2)$$
$$\text{Recall} = TP / (TP+FN) \quad (3)$$
$$\text{F1 score} = 2*(\text{Precision}*\text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

TP, TN, FN, FP represents the number of true positive, true negative, false negative, false positive respectively.

Models	**Accuracy**	**Precision**	**Recall**	**F1 Score**
Inception V3	89.99	90.57	93.58	91.01
ResNet50	90.06	92.48	91.53	92.01
VGG16	83.49	93.35	79.23	85.71
Dense Net	85.67	84.98	85.55	85.26
Custom CNN	92.7	95.03	93.33	94.1

Table 2 Results of networks by Accuracy, Precision, Recall and F1 Score

VII.RESULTS

Corresponding to these five CNN models, we plotted the training set accuracy, training set loss, validation set accuracy (Val accuracy), and validation set loss (Val loss) in four-line charts for comparison (as illustrated in Figures 3-6). In this situation, setting epoch to 50 allows you to compare the accuracy of different algorithms over the same number of iterations, so that the speed of training between algorithms can be reflected. If you come across a new type of pneumonia and time is of the essence, researchers must train the lung images of the new type of pneumonia as soon as possible, so achieving a higher accuracy rate while saving time is also a factor to consider. Following that, we applied the five trained models to the experiment set and measured their accuracy and recall. The results show that the accuracy of pneumonia recognition using Custom CNN can reach 92.79%, with a recall of 93.33%. Furthermore, Figures (3-6) show that the Custom CNN has higher accuracy and lower loss while using fewer floating-point calculation. In comparison to other networks, Custom CNN can trade higher data throughput for lower accuracy.

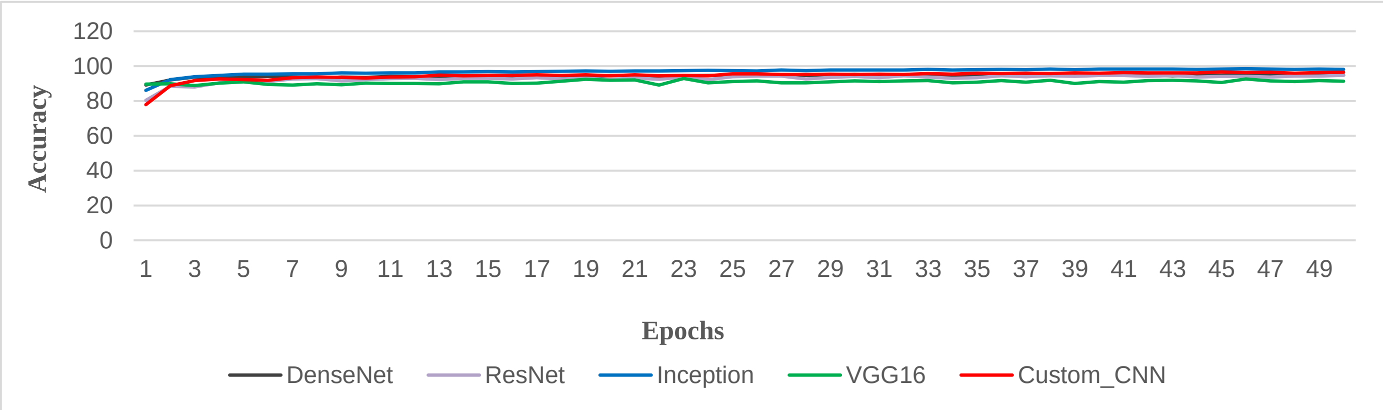

Figure 3 Training set Accuracy of five CNNs

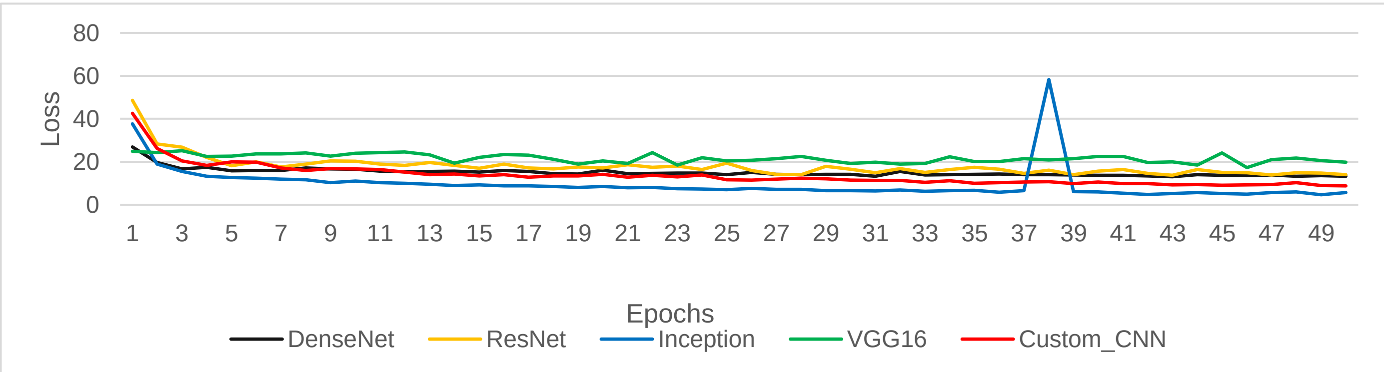

Figure 4 Training set Loss of five CNNs

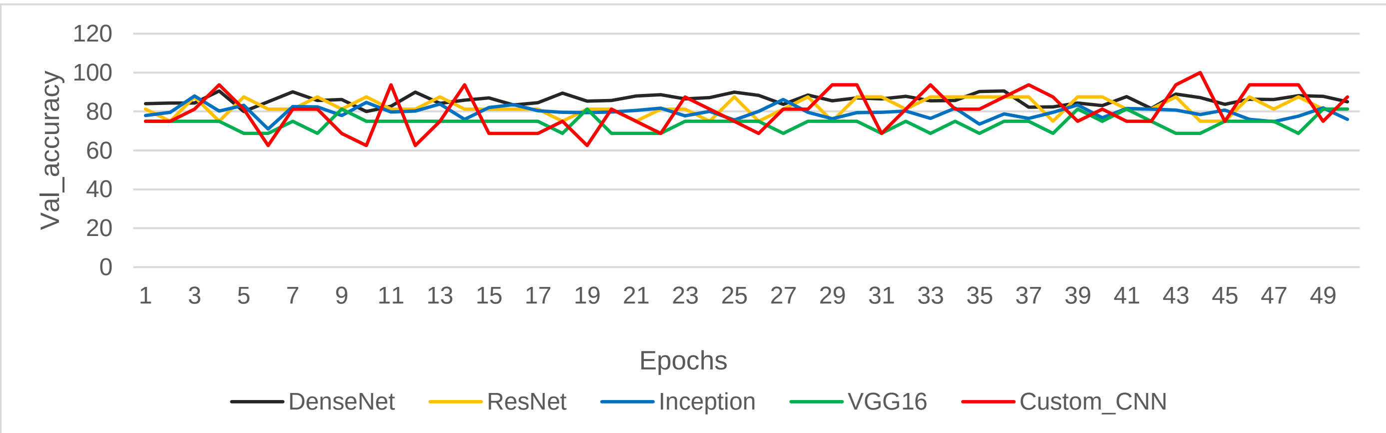

Figure 5 Validation set Accuracy of five CNNs

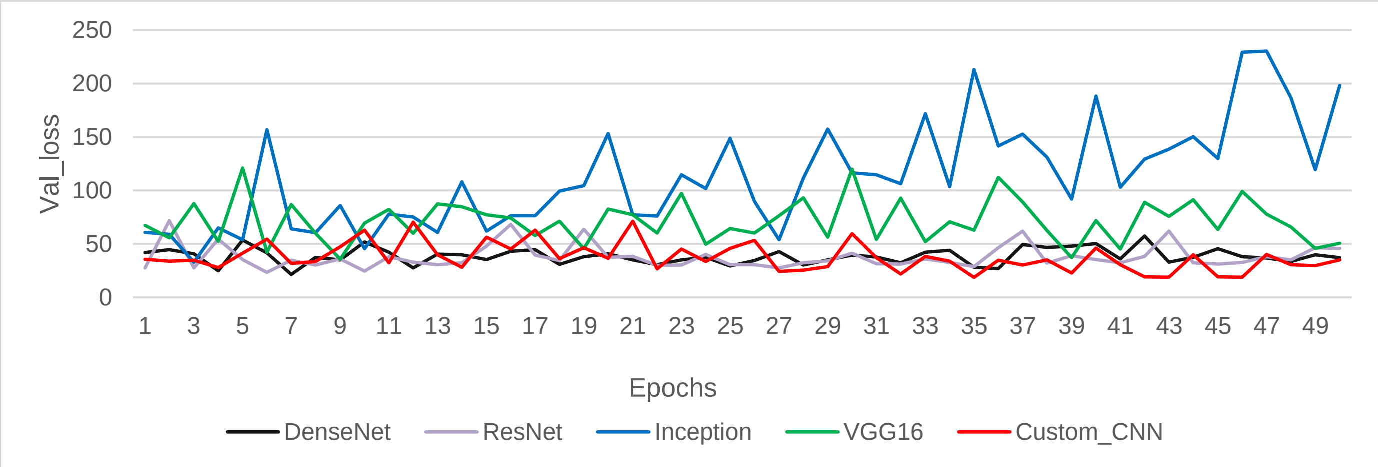

Figure 6 Validation set Loss of five CNNs

VIII. CONCLUSION

Thus, this study paper compares the accuracy of five used deep learning models used to diagnose clinical data on a dataset consisting of X-ray images of the lungs with pneumonia and normal lungs. Because of Custom CNN 's superior performance, we focus on its network structure. The results show that all five network structures can recognize pneumonia, and Custom CNN's accuracy is higher than that of other network structures. The custom CNN a 5-layer network gave about 92.7% accuracy compared to other models. Furthermore, the use of artificial intelligence technology in the medical field is insufficient, and the dataset in this field needs to be improved in terms of type. The performance of CNN-based pneumonia diagnosis algorithms will improve as the amount of pneumonia image data increases and the network structure improves. In the future, the use of clinical image diagnosis of pneumonia X-rays may reduce clinician workload and allow patients to receive early diagnosis and treatment, lowering pneumonia mortality rates.

REFERENCES :

1. 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. 202
2. R. Siddiqi, “Automated Pneumonia Diagnosis using a Customized Sequential Convolutional Neural Network,” Proceedings of the 2019 3rd International Conference on Deep Learning Technologies - ICDLT 2019, 2019
3. S. Yao, Y. Chen, X. Tian, and R. Jiang, “Pneumonia Detection Using an Improved Algorithm Based on Faster R-CNN,” Computational and Mathematical Methods in Medicine, vol. 2021, p. 8854892, 2021
4. T. Gabruseva, D. Poplavskiy, and A. A. Kalinin, “Deep Learning for Automatic Pneumonia Detection,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1436–1443, Jun. 2020
5. S. L. K. Yee and W. J. K. Raymond, “Pneumonia Diagnosis Using Chest X-ray Images and Machine Learning,” Proceedings of the 2020 10th International Conference on Biomedical Engineering and Technology, Jun. 2020.
6. A. Mabrouk, R. P. Díaz Redondo, A. Dahou, M. Abd Elaziz, and M. Kayed, “Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks,” Applied Sciences, vol. 12, no. 13, p. 6448, Jan. 2022.
7. G. Wang et al., “A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images,” Nature biomedical engineering, vol. 5, no. 6, pp. 509–521, Jun. 2021.
8. N. M. ud din, R. A. Dar, M. Rasool, and A. Assad, “Breast cancer detection using deep learning: Datasets, methods, and challenges ahead,” Computers in Biology and Medicine, vol. 149, p. 106073, Oct. 2022.
9. A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, “Identifying pneumonia in chest X-rays: A deep learning approach,” Measurement, vol. 145, pp. 511–518, Oct. 2019.
10. G. Liang and L. Zheng, “A transfer learning method with deep residual network for pediatric pneumonia diagnosis,” Computer Methods and Programs in Biomedicine, vol. 187, p. 104964, Apr. 2020.
11. World Health Organization, “Pneumonia,” Who.int, Nov. 11, 2021.
12. M. I. Neuman et al., “Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children,” Journal of Hospital Medicine, vol. 7, no. 4, pp. 294–298, Oct. 2011.
13. O. Stephen, M. Sain, U. J. Maduh, and D.-U. Jeong, “An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare,” Journal of Healthcare Engineering, Mar. 27, 2019.
14. P. Rajpurkar et al., “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,” arXiv.org, 2017.
15. A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, “Pneumonia Classification Using Deep Learning from Chest X-ray Images During COVID-19,” Cognitive Computation, Jan. 2021.