International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Deep Learning Approach for Pneumonia Detection from X−ray Images

B. R. Kanawade*[1], S. N. Zaware[2], Janvi Nandre[3], Yashashree Mahale[4], Khushi Dhake[5]

Submitted: 12/11/2022 **Accepted:** 15/02/2023

Abstract: Pneumonia, which is caused by Streptococcus Pneumoniae, can be deadly if undetected or mistreated. The most common approach for detecting Pneumonia is to have a professional radiologist review a chest X-ray picture, which takes longer and is less reliable. Professionals and physicians can employ computer-assisted diagnosis to solve this problem. Computer-assisted diagnosis might improve doctors' ability to make quick and accurate judgments. Convolutional neural networks that are abbreviated as CNNs have become particularly popular in disease classification due to the usefulness of algorithms in deep learning for the analysis of medical images. The performance of some pre-trained CNN models was examined to reach the final result, followed by an ensemble of top-performing models. The study revealed that putting together various pre-trained CNN models can improve detection accuracy, with the best accuracy being 94.39%.

Keywords: *Pneumonia Detection, Computer-aided diagnosis, Convolutional Neural Network, Ensemble.*

1. Introduction

Pneumonia, commonly known as Streptococcus Pneumoniae, is a life-threatening condition that develops quickly. Coughing up green, yellow, or red mucus is the most common symptom of Pneumonia. Pneumonia is the top cause of death in children throughout the world, with India accounting for 20% of those fatalities and having the highest burden of pediatric Pneumonia of any nation. Most of the children under the age of five lost their life due to pneumonia and the count in 2018 for the same was nearly 800,000. Despite the fact that treating Pneumonia in its whole is crucial for infant survival, progress has been slow. Radiologists react to white spots on X-ray pictures to indicate Pneumonia, however owing to its restricted color range, it is difficult to diagnose. This affirmation can only be done in the later phase of the sickness. Hence, early recognition of the disease isn't applicable using manual techniques and the overlapping organs with blurred boundaries, make it difficult for radiologists to diagnose Pneumonia through chest radiography.

In recent years, medical image processing has become a research focus in the field of computer vision. Due to the time-consuming process, researchers have developed several computer algorithms and computer-aided diagnostic tools to analyze x-ray images. They have not, however, shown to be especially beneficial in assisting professionals in making decisions. Deep Learning techniques have recently been used by medical imaging researchers to review and categorize medical pictures for the identification of various diseases such as brain tumors, skin cancer, breast cancer, and TB. This has prompted us to present our work on a deep learning strategy for detecting Pneumonia in the chest using X-ray pictures.

Artificial Neural Networks (ANNs) are algorithms encouraged by the composition and operation of the human brain, and Deep Learning is an area of Machine Learning that deals with them. CNN is the most extensively utilized neural network architecture in the field of medical imaging in Deep Learning. CNN is a sort of deep neural network that conducts convolution computations and has a linked topology. CNN also has a feature which is a learning feature that can capture spatial local correlations of input data using convolution operations to capture transformation invariance. CNN has the benefit of not requiring complicated picture pre-processing and being able to use original images as direct input. As a result, they may be used in a variety of situations.

In this article, the (Visual Geometry Group) VGG16, VGG19, Densenet121, Densenet201, Densenet169, Mobilenet, mobilenetV2, InceptionResnetV2, and Xception convolutional neural network variations were investigated. With a small quantity of data, a greater detection precision was achieved using the ensemble approach on top-

1,3,4,5 Department of Information Technology
International Institute of Information Technology,
Pune, India
2Computer Engineering Department,
AISSMS Institute of Information Technology, Pune, India
1kanawadebr@gmail.com,2sarika.zaware@aissmsioit.org
3nandrejanvi@gmail.com,4yashashree125@gmail.com,5khushid hake7@gmail.com

performing models. The maximum detection accuracy was attained, with a score of 94.39 percent.

2. Model Variants

Multiple variants of convolutional neural networks were tested namely VGG16, VGG19, DenseNet201, DenseNet121, DenseNet169, Xception, MobileNet, MobileNetV2 and InceptionResnetV2. VGG16, Xception, and DenseNet201 outperformed all other selected models.

Table 1. Pre-trained Model Variants for Image Classification

Model	**Image Size**	**Number of Layers**
VGG16	224 * 224 * 3	16
VGG19	224 * 224 * 3	19
MobileNet	224 * 224 * 3	28
MobileNetV2	224 * 224 * 3	53
Xception	299 * 299 * 3	36
Inception ResNet V2	299 * 299 * 3	164
DenseNet201	224 * 224 * 3	201
DenseNet121	224 * 224 * 3	121
DenseNet169	224 * 224 * 3	169

VGG16 (also known as OxfordNet) is a convolutional neural network architecture named after the Oxford-based Visual Geometry Group. It has 3x3 filtered convolution layers with stride 1 and uses the same padding, as well as a maxpool layer which is of 2x2 filter of stride 2 for the entire model. This architecture allows VGG16 to outperform baselines on many tasks and datasets outside of imageNet.

The DenseNet201 is a 201-layer convolutional neural network. In comparison to other models, it proposes a more radical dense connection mechanism. Each level or layer of DenseNet receives additional input from all previous levels and sends its own feature map to all subsequent levels. Characteristic reuse is another important feature of DenseNet, which is accomplished by connecting features on channels. With fewer parameters and lower computational costs, DenseNet is able to achieve better performance.

Xception is a convolutional neural network that is 71 layers deep and Inspired by Google's Inception model.The architecture of Xception is explained as a linear stack of depthwise separable convolution layers with residual connections which makes mapping of both cross-channels correlations and the spatial correlations entirely decoupled and reduces complexity significantly as compared to other normal convolutional operations.

3. Related Work

Several strategies for detecting Pneumonia disease using chest X-ray images have been published in the literature. Some methods use a feature extraction approach in combination with machine learning algorithms to classify chest x-rays, while others use a deep learning approach to feature extraction and classification.

Without the help of a professional radiologist, interpreting chest X-rays can be difficult. Several researchers are working to automate X-ray processing. To overcome limited color space in grayscale X-ray images, Deniz et al. [1] used various preprocessing techniques like increasing the contrast and lightening the images artificially. To Remove the noise or distortion present in chest X-ray images, which eventually leads to incorrect results, Min-jen-Tsai et al. [2] used adaptive mean filtering techniques. In image processing, adaptive filters are widely used to enhance or restore data by reducing noise without dramatically obscuring the image's structures. To remove the image regions which is found to be noise corrupted Adaptive mean filter is used.

Huaiguang Wu et al. [3] investigated how pre-processing affects model accuracy, finding that different input image sizes yield diverse outcomes. Data imbalance leads to poor performance and the classification may be biased. Okeke Stephen et al. [4], Garima et al [5] and Tahsifer Rahman et al. [6] used data augmentation techniques to work on data imbalance problems. Data expansion allows practitioners to significantly expand the variety of data available in their training models without actually collecting additional data. M.Togaçarv [7] used mRMR technique for feature extraction. In this work, accuracy is enhanced using mRMR. To enhance the identification of Pneumonia lesions Shangjie et al. [8] used an attention mechanism and dilated convolution along with a Yolo. Double K-means were used to locate the area and also improve localization accuracy. Amit Kumar Jaiswal et al. [9] described an algorithm that can detect the visual signal for Pneumonia in medical chest X-rays also called radiographs, and provide an output as to whether it is Pneumonia positive or negative. Additionally, if the signal indicates a positive result, the predicted bounding boxes around the lung opacities is also returned. For this, the authors used Mask- R CNN which was used to identify the lung opacity which is likely to depict the Pneumonia disease.

Ilyas Sirazitdinova et al. [10] used an approach that was based on an ensemble of Mask-R CNN and RetinaNet . Pneumonia is a lung infection that affects a small area of the chest whereas X-ray is a daunting task for modern object detectors. This is a major issue and to tackle this issue authors used the FPN (Feature Pyramid Network) concept as the backbone of both models to solve this problem since FPN produces multi-scale feature maps with higher quality details than that of the default feature pyramid. The authors emphasized the advantages of focal loss and object-detection approaches in the terms of classification metrics. Dimpy et al. [11] used the process of transfer learning and fine-tuning method to detect Pneumonia. A comparison of all these CNN models was done using accuracy as a metric which shows noteworthy results

4. Proposed System

4.1. Outline of the methodology

Proposed System is summarized in Figure 1. Preprocessing of chest X-ray images, data augmentation, transfer learning using VGG19, DenseNet121, Xception neural networks, and ensemble categorization are all part of the process. The steps are then detailed in greater depth in the following subsections.

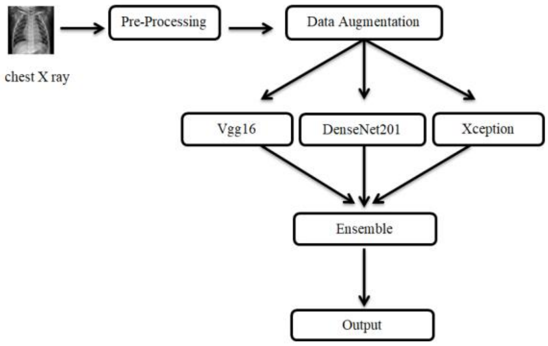

Fig. 1 Pneumonia Detection System

4.2. Dataset

In this study we used the dataset of 5856 frontal chest X-ray jpeg images from Kaggle. The image size in the collection ranges in resolution from 712x439 to 2338x2025 pixels. The collection contains 1583 photos of normal cases and 4273 photos of Pneumonia cases. In our models, 0 indicates normal class and 1 indicates Pneumonia class.

Table 2. Kaggle Chest X-Ray Dataset for Pneumonia Detection

Table Head	Train	Validation	Test
Pneumonia	3106	777	234
Normal	1079	270	624
Total	4185	1047	390

4.3. Preprocessing and Data Augmentation

The above training dataset was split into 80:20 ratios with 80% being used for training deep learning models and 20% for validation of trained models. table 2 represents the distribution of the dataset.

Chest X-ray images in the dataset are of different resolutions. The CNN models used in the study also require images of the same size for training as well as validation and testing. So we have used 300 x 300 resolution for our study. Another way of decreasing the computational cost is to normalize values of pixels in the range from 0-255 to 0-1. The purpose of normalizing is to transform the values of numeric features to a common scale without distorting feature range or erasing information. Normalization is a scaling technique that shifts and rescales the values to make them range between the values 0 and 1.

A huge amount of data is required for deep learning to acquire better results. However, there may not be enough data for deep learning models, especially in the medical field. It is a very expensive and time-consuming process to obtain this data. There is also an issue of data imbalance with the available dataset which could lead to biased results. To overcome these issues various data augmentation techniques can be used to avoid overfitting and get reliable results.

Data augmentation strategies shown in figure 2 are methods used for modifying training data in such a way that the representation of the array is changed while the labeling remains the same. The other augmentations used in this study are Grayscales, horizontal and vertical flips, random crops, color jitters, translations, rotations and many more. We can easily improve our training data by performing a few of these adjustments. By applying only few adjustments to our training data we can easily double or quadruple the number of training examples which can develop a very robust model.

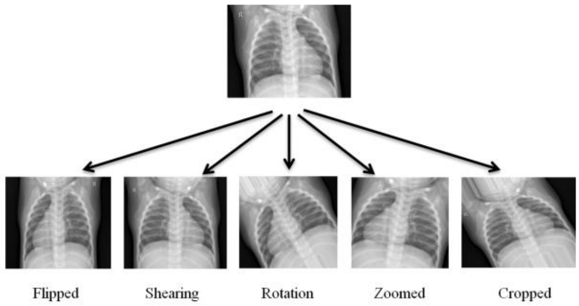

Fig. 2 Data Augmentation

4.4. Convolutional neural network and variant

A Convolutional Neural Network (CNN) is a Deep Learning algorithm that takes an input image and assigns weights which are learnable to various attributes in the image. When it is compared to other classification methods, the amount of preprocessing required by a CNN is remarkably less. While filters are hand-engineered in primitive methods, CNN can learn these filters/characteristics with proper training on its own. The architecture of CNN is inspired by the layout of the visual cortex and resembles the connection pattern of neurons in the human brain.

The convolutional neural network (CNN) is formed by an input layer, hidden layers, and an output layer. Any intermediate layers of feedforward neural networks are said to be hidden because the masking of input and output is done by activation function and final convolution. The hidden layer of the convolutional neural network performs the convolution. In the convolution, the layer of the dot product of the convolution kernel and layer input matrix is included. Relu is the activation function of this product which is the

Frobenius inner product. The convolution kernel slides along the input matrix of one layer, contributing to the next layer's input, and generating a feature map. After that, other layers such as pooling layers, fully connected layers, and normalized layers are added.

The pooling layer reduces the size of convolved features spatial. Dimensionality reduction cuts down on the amount of computing power needed to process the data. It is also beneficial for extracting rotational and positional invariant dominant features, which helps keep the model's training process running smoothly.

Table 3 shows the various CNN-based pre-trained models used in this Study. Figure 4 shows the comparative results of these models

The models used are like VGG16, VGG19, DenseNet121, DenseNet201, DenseNet169, InceptionResnet, MobileNet, MobileNetV2, and Xception.

For the training of all the models, 300 X 300-pixel chest X-ray images were used. All the models were trained on 35 epochs. The outcomes for three tries are shown in the table below.

According to the results, Xception, VGG16, and DenseNet201 outperformed the other models. The highest accuracy for models Xception, VGG16, and DenseNet201 after three attempts was 92.78%, 93.26%, and 93.26% respectively.

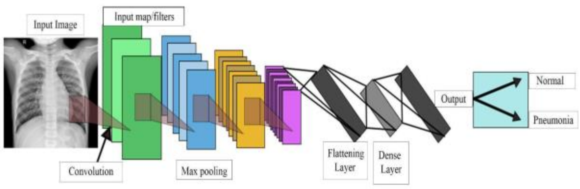

Fig. 3 Convolutional Neural Network

Table 3. Model Variant Results

Model	35 epochs 300 X 300 - trial 1	35 epochs 300 X 300 - trial 2	35 epochs 300 X 300 - trial 3	Best Results
VGG16	93.26	91.34	91.66	93.26
DenseNet201	93.26	92.14	93.03	93.26
Xception	87.98	89.58	92.78	92.78
DenseNet121	92.14	92.62	91.80	92.62
MobileNetV2	87.66	92.14	91.66	92.14
MobileNet	91.60	88.3	89.42	91.60
DenseNet169	91.50	91.50	90.06	91.5
Inception ResNetV2	74.8	87.98	88.94	88.94
VGG16	89.74	89.26	86.21	89.74

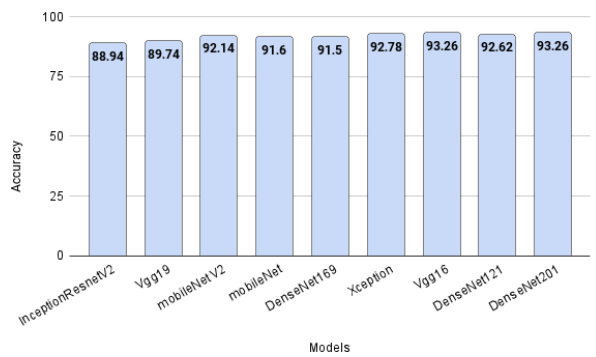

Fig. 4 Different Model Variant Result Comparison

4.5. Ensemble Learning

Ensemble learning is a machine learning methodology that combines multiple base models to build a single best-predictive model. Ensemble learning can be a very powerful yet simple tool for increasing the final results.

After the individual results from various pre-trained models were received, we ensembled the top 3 models which outperformed the remaining models, which were VGG16, Xception, and DenseNet201 to get the final result. The ensemble technique gives improved accuracy. The ensemble classification strategy shown in figure 5 was employed to integrate the predictions of three pre-trained neural networks. To arrive at a final prediction, the outputs of pre-trained neural networks were combined into a prediction vector, and majority voting was used to reach the final conclusion.

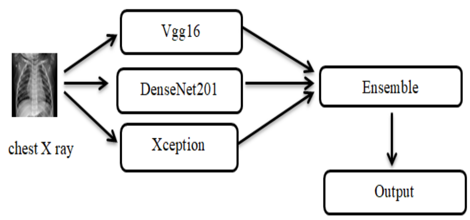

Fig. 5. Ensemble Learning

5. Experimental Results

In this paper, various variants of pre-trained CNN models were tested like VGG16, VGG19, DenseNet121, DenseNet201, DenseNet169, InceptionResnet, mobileNet, mobileNetV2, and Xception. According to the results, Xception, VGG16, and DenseNet201 outperformed the other models. Hence these top-performing models were used to obtain the final results by ensembling. The overall accuracy of our system achieved after ensembling of VGG16, Xception, and Densenet201 is 94.39%. Table 4 depicts the detailed results of our system for each class.

Table 4 Proposed System Results

624	Precision	Recall	F1-Score	Support
Normal	95.43	89.31	92.26	234
Pneumonia	93.82	97.43	95.59	390
Accuracy			94.39	624
Total	94.63	93.37	93.93	624

6. Conclusion

In this study, the goal was to propose a Deep Learning-based approach to classify Pneumonia from chest X-ray images which could help doctors make better decisions. Compared to the work done in this field, the proposed system displays better performance; even with a lesser dataset due to the adoption of various techniques like data augmentation and ensemble approach. As a future enhancement, the proposed system can be tested on larger datasets which could in turn help the system give enhanced results as compared to the existing system.

Author contributions

Conceptualization, B.R.K. and J.N.; methodology, B.R.K; methodology, B.R.K., S.Z., J.N., Y.M and K.D.; investigation, B.R.K., J.N., Y.M.. and K.D.; formal analysis, B.R.K, S.Z., and K.D; writing—original draft preparation, J.N., K.D. and Y.M.; writing—review and editing, J.N. and K.D.; project administration, B.R.K;

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Saul, C. J., Urey, D. Y., & Taktakoglu, C. D. (2019). Early diagnosis of pneumonia with deep learning. arXiv preprint arXiv:1904.00937.

[2] Tsai, M. J., & Tao, Y. H. (2019, December). Machine Learning Based Common Radiologist-Level Pneumonia Detection on Chest X-rays. In 2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS) (pp. 1-7). IEEE. DOI: 10.1109/ICSPCS47537.2019.9008684

[3] Wu, H., Xie, P., Zhang, H., Li, D., & Cheng, M. (2020). Predict pneumonia with chest X-ray images based on convolutional deep neural learning networks. Journal of Intelligent & Fuzzy Systems, (Preprint), 1-15. DOI: 10.3233/JIFS-191438

[4] Stephen, O., Sain, M., Maduh, U. J., & Jeong, D. U. (2019). An efficient deep learning approach to pneumonia classification in healthcare. Journal of healthcare engineering, 2019. DOI : https://doi.org/10.1155/2019/4180949

[5] Verma, G., & Prakash, S. (2020). Pneumonia classification using deep learning in healthcare. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(4), 1715-1723.

[6] Rahman, T., Chowdhury, M. E., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., ... & Kashem, S. (2020). Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Applied Sciences, 10(9), 3233. DOI: https://doi.org/10.3390/app10093233

[7] Toğaçar, M., Ergen, B., Cömert, Z., & Özyurt, F. (2020). A deep feature learning model for pneumonia detection applying a combination of mRMR feature selection and machine learning models. Irbm, 41(4), 212-222. DOI: https://doi.org/10.1016/j.irbm.2019.10.006

[8] Yao, S., Chen, Y., Tian, X., Jiang, R., & Ma, S. (2020). An Improved Algorithm for Detecting Pneumonia Based on YOLOv3. Applied Sciences, 10(5), 1818. DOI: https://doi.org/10.3390/app10051818

[9] Jaiswal, A. K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J. J. (2019). Identifying pneumonia in chest X-rays: A deep learning approach. Measurement, 145, 511-518. DOI: https://doi.org/10.1016/j.measurement.2019.05.076

[10] Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R., & Ibragimov, B. (2019). Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. Computers & Electrical Engineering, 78, 388-399. DOI: https://doi.org/10.1016/j.compeleceng.2019.08.004

[11] Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R., & Mittal, A. (2019, February). Pneumonia detection using CNN based feature extraction. In 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) (pp. 1-7). IEEE. DOI: 10.1109/ICECCT.2019.8869364