Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

A Comparative Study on the Detection of Pneumonia in Chest X-Ray Images using Deep Learning Models

Remya B Nair, Dr. Cini Kurian

Computer Science, Al-Ameen College, Edathala, India

This paper offers a comparative analysis between the VGG16 architecture with Support Vector Machine (SVM) classifier and a custom Convolutional Neural Network (CNN) for detecting pneumonia from chest X-ray images. Utilizing a publicly accessible dataset, both models underwent training and evaluation based on several metrics, such as accuracy, precision, recall, and F1 score. The results highlight the benefits and drawbacks of each approach, providing insight into the effectiveness and efficiency of using VGG16 in contrast to CNN architectures for medical image analysis.

1. Introduction

Pneumonia, a serious and infectious disease, affects millions of individuals. Most pneumonia patients are elderly and have chronic illnesses like asthma or diabetes. When diagnosing pneumonia, chest X-rays are thought to be the most accurate method of identifying the extent and location of the contaminated zone in the lungs. According to pulmonologists, X-ray imaging is an essential investigation for patients with suspected lung problems. X-ray scans provide a clear view of the internal architecture of the lungs, allowing for the identification of abnormalities or diseases. They are particularly adept in spotting lung-related conditions like tumors, TB, pneumonia, and others. Deep learning-based artificial intelligence (AI) techniques have become more and more common in the medical imaging industry, particularly for X- ray.

This study compares the CNN architecture and VGG16, a popular pre-trained convolutional neural network, for identifying pneumonia from chest X-ray images. Several performance indicators, such as accuracy, precision, recall, F1-score, and computing efficiency, are used in the study to assess both models. A large collection of labelled chest X-rays is used to train and test the models, guaranteeing accurate and consistent comparisons. The results of this study focus on the advantages and disadvantages of both architectures, which aid in the creation of optimal deep-learning solutions for clinical applications in the diagnosis of pneumonia.

2. Related work

Numerous experiments have been carried out to use CNN to detect pneumonia from chest X-ray pictures. VGG16 and other pre-trained models have shown efficacy because of their strong feature extraction capabilities. Custom CNN architectures, on the other hand, provide more design flexibility and might perform better on specific activities. Although there have been several comparison studies, our work evaluates VGG16 and a customised CNN for detecting the presence of Pneumonia in X-Ray Images.

In this section, an overview of current research activities on deep learning-based models for the analysis of X-ray images are given. This study reviewed various advancements in the data preprocessing techniques, model construction, and model performance for the analysis of Pneumonia detection.

To differentiate between COVID-19 and pneumonia patients from chest X-ray pictures, Musha et al. created a CNN-based algorithm [1]. They underlined how crucial data augmentation and preprocessing are to improving the model's functionality. The model showed great accuracy, precision, and recall using a variety of publicly available datasets, demonstrating its potential for quick diagnosis. The scientists did, however, acknowledge that the model's generalisability across a range of patient populations was limited, and they urged more practical testing

Yang et al. introduced a deep learning framework using Explainable AI (XAI) for pneumonia detection [2]. This work aimed to improve the deep learning model' interpretability by considering the background features of the image. To enhance clinicians', trust their model provided valuable insights. The study showed that the background features improved the model's diagnostic accuracy. Besides these advancements this research identified challenges in managing background complexity and suggested the refining of the model for broader clinical applicability.

Barhoom and Abu Naser proposed the VGG16 architecture for pneumonia classification [3]. In their model extensive data augmentation techniques were implemented to optimize the model's performance. They resized the images to 128x128 pixels to balance computational efficiency. They also use normalization for consistency. The VGG16 model exhibited strong classification results, but its computational demands could hinder deployment in resource-limited settings. This study demonstrated the model's efficacy and highlighting the complexity of the architecture.

Malik et al. proposed CDC Net, a CNN model designed for multi-class disease detection [4]. Their model analysed the X-ray images and classified as pneumonia, COVID 19 and other chest conditions. Into their model they integrated dilated convolutions and residual networks to improve the feature extraction process. CDC Net achieved an impressive area under the curve (AUC) of 0.9953. This model outperforming traditional models like VGG-19 and ResNet-50. The study underscored the potential of deep learning for simultaneous detection of multiple diseases. But also emphasized the need for further validation to ensure clinical reliability.

A machine learning approach was proposed by Barakat et al. for paediatric pneumonia detection [5]. They emphasized the benefits of machine learning models over deep learning in terms of interpretability and computational efficiency. Their model used data augmentation to address class imbalance. They used the Quadratic SVM for classification and achieved an accuracy of 97.58%, significantly reducing classification time compared to transfer learning models. The authors highlighted the need for larger datasets to improve model robustness.

3. METHODOLOGY

A. Dataset Description

We used the publicly available Kaggle dataset for this study. It consists of chest X-ray images labelled Pneumonia and Normal. The dataset was pre-processed and split into training, validation, and test sets to ensure effective training of the model and evaluation [8].

B. Data Collection

Images were sourced from a publicly available repository, which provided a diverse set of X-ray images for both pneumonia and non-pneumonia cases[8]. A sample X-ray image of pneumonia and Normal image is shown in Fig1 and Fig 2.

Fig.1 Image depicts Pneumonia affected Chest X-ray image

Fig.2 Image depicts Normal Chest X-ray image

Model Architectures

This section deals with the detailed description of the applied methodology. First discuss with Convolution Neural Network followed by VGG 16 model[6 -7]. Block diagram of the proposed model is shown in Fig.3.

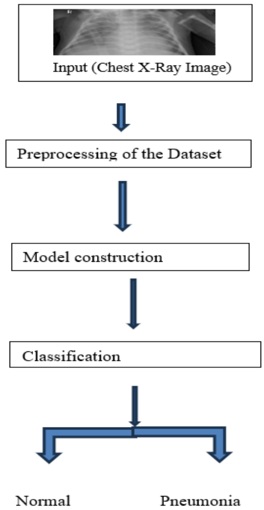

Fig 3. Block diagram of the proposed model

C. Convolution Neural Network Architecure (CNN)

A sequential neural network in the proposed model, where the output of one layer is provided as input to the following layer. Sequential neural network provides an easy-to-use interface for building simple neural networks, where layers are added one after another in a linear fashion, The model consists of the following steps.

a. Preprocessing

Prior to feeding the images into the model, the following preprocessing steps are applied

- Step 1: Resizing

Images are resized to 224x224 pixels to match the input size required by the CNN.

- Step 2: Normalization

Pixel values are normalized to the range [0, 1] by dividing by 255, which helps in speeding up training and stabilizing the gradient descent.

- Step 3: Data Augmentation

Techniques such as rotation, flipping, and zooming are applied to increase the diversity of the training dataset and reduce overfitting.

- Step 4: Splitting

The dataset is split into 70% for training, 15% for validation, and 15% for testing.

b. CNN Architecture

The model follows a standard deep CNN architecture comprising several convolutional and pooling layers, followed by fully connected layers. Below is the detailed description of each layer:

a. Convolutional Layers

It is the model's conventional deep neural network architecture consists of convolutional and pooling layers, followed by fully connected layers. Each layer is explained further below.

a. Convolution Layer

It acts as CNN's core component. The convolution technique is used in mathematics to mix two functions. The input image is initially transformed into a matrix for the CNN models. The input matrix is passed through a convolution filter, which multiplies the elements and saves the result. A feature map is generated as a result. When black-and white images are taken, the 3D filter is typically used to produce 2D feature maps. When the input image is represented as a three-dimensional matrix, convolutions are performed in three dimensions, with the third dimension indicated by the RGB color.

The model starts with many 3x3 filter 2D convolutional layers, which extract the input image's spatial features. After the first three convolutional layers, each employing 64 filters, ReLU activation functions are used to add nonlinearity.

The following two convolutional layers, which employ 128 filters apiece to capture increasingly intricate patterns, deepen the feature maps.

b. Max-Pooling Layers

After each set of convolutional layers, max-pooling operations (with a pool size of 2x2) are applied. This helps in reducing the spatial dimensions of the feature maps and makes the network more computationally efficient, while retaining the most important features.

c. Flatten Layer

The output of the last convolutional layer (28x28x128) is flattened into a 1D vector to get it ready for the fully connected layers.

d. Fully Connected Layers (Dense Layers)

The first fully connected layer consists of 100 units and uses ReLU activation. The second fully connected layer consists of 200 units, also with ReLU activation, to learn complex combinations of features. A total of 10,056,103 parameters are generated from these three dense layers The final layer is a softmax layer with 3 units, representing the classes: pneumonia, and Normal

e. Model Training

Following parameters are used for model training.

- Loss Function

The model uses categorical cross-entropy as the loss function, which is suitable for multi-class classification problems.

- Optimizer

The Adam (Adaptive Moment Estimation) optimizer is employed for training, as it adapts the learning rate during training, improving convergence and reducing the risk of overshooting minima.

- Batch Size and Epochs

The model is trained using a batch size of 32 and 50 epochs, with early stopping implemented to prevent overfitting. The training process is monitored by tracking the validation loss and accuracy.

D. Model Architecture Design(Vgg16)

We employed a pre-trained VGG16 model, for feature extraction. For this fine-tuning the top layers to adapt to the pneumonia detection task. The output layer was modified to a single neuron with a sigmoid activation function for binary classification. This VGG 16model construction includes preprocessing, feature extraction and Classification using Support Vector Machine (SVM)

a. Dataset Preprocessing

The dataset used consists of chest X-ray images. The following preprocessing steps were applied to ensure consistency and compatibility with the deep learning model:

- Step 1: Resizing

All images were resized to a uniform dimension of 224×224 pixels. This size was chosen to align with the input requirements of the pre-trained VGG16 model.

- Step 2: Grayscale to RGB Conversion

If any images were grayscale, they were converted to RGB format using OpenCV's cv2.cvtColor function. This step ensures compatibility with the VGG16 model, which requires three input channels.

b. Feature Extraction Using Vgg16

A pre-trained VGG16 model, available from the ImageNet library, was utilized to extract deep features from the chest X-ray images. The steps are as follows:

- Step 1: Model Configuration

Weights that had previously been trained on ImageNet were fed into the VGG16 model.

- Step 2: Feature Extraction

The pre-processed images have been processed through the model to extract features. A 3D feature map was generated by using the output of the last convolutional layer of VGG16 as the feature representation for every image.

- Step 3: Feature Flattening

The feature maps were flattened into 1D feature vectors to get the extracted features ready for classification. In this stage, the feature space's dimensionality is decreased while crucial classification information is preserved

- Step 4: Data Splitting

The dataset was divided into subsets for testing and training: This is how the split ratio looks.20% of the data was set aside for testing, while the remaining 80% was used for training. The features were divided with the matching labels.

c. Classification Using Support Vector Machine (SVM)

Pneumonia and non-pneumonia classifications were created using a Support Vector Machine (SVM). Here we are using a radial basis function (RBF) kernel is used to map the features [13]. The flattened feature vectors and the labels were used to train the model. The trained SVM model was used to make predictions on the test dataset.

d. Evaluation Metrics

The performance [7] of the classifier was evaluated using the following metrics:

- Accuracy

Calculated as the proportion of correctly classified samples out of the total number of samples.

- Classification Report

Detailed metrics including precision, recall and F1-score were generated for each class [12].

4. RESULTS AND DISCUSSION

The performance of the VGG16 and CNN models for pneumonia detection is compared based on standard evaluation metrics [12] such as Accuracy, Precision, Recall and F1 Score. Below is a detailed analysis of the results

- Accuracy

The model is evaluated on the test dataset to measure the accuracy of the model, which is computed by comparing the predicted labels from the model with the ground truth labels from the dataset [12].

The formula for accuracy calculation is as follows

$$\text{Accuracy} = P/N*100 \qquad (1)$$

P indicates Number of Correctly Predicted Samples, and N indicates Total Number of Samples in equation (1)

It can also be noted as P=TP+TN and N=TP+TN+FP+FN. Here TP indicates True positives, TN is true negatives, FP is false positives and FN is false negatives [9].

a. Recall

It measures the ability of the model to correctly identify positive samples (true positives) out of all the actual positive samples (true positives + false negatives). In other words, recall [12] assesses the model's ability to find all relevant instances of a particular class.

$$\text{Recall} = TP/(TP+FN) \qquad (2)$$

In Equation (2) TP is true positives and FN is false negatives [9].

A high recall indicates that the model is effectively identifying most of the positive samples, while a low recall suggests that the model is missing a significant number of positive samples.

b. Precision

It represents the ratio of the number of correctly predicted positive samples to the total number of predicted positive samples [12].

Precision =TP/((TP+FP) (3)

In Equation (3) TP is True Positives and FP is False Positives

c. F1 Score

It provides a balanced measure of both precision and recall. The F1 score [12] considers both false positives and false negatives and provides a single value that combines the two metrics shown in equation 4.

F1 Score = 2 * (Precision * Recall) / (Precision + Recall) (4)

Performance Metrics of the CNN with SVM and VGG 16 models are listed in the Table1 and Table 2.

Table 1 results obtained by VGG 16 with svm classifier

Metric	Value
Accuracy	0.9459
Precision	0.9479
Recall	0.9474
F1 Score	0.9486

Table 2 Performance metric of CNN

Metric	Value
Accuracy	0.87
Precision	0.88
Recall	0.8979
F1 Score	0.8886

According to these findings, both models are effective at identifying pneumonia, but the VGG16 model is a better option for applications that demand high specificity and precision, like automated diagnosis systems in advanced medical settings. The CNN model, on the other hand, might be better suited for quick implementations in settings with limited computational resources because to its simpler structure and comparatively lower performance. The generalisability and interpretability of these findings must be validated by more testing on more extensive and diverse datasets.

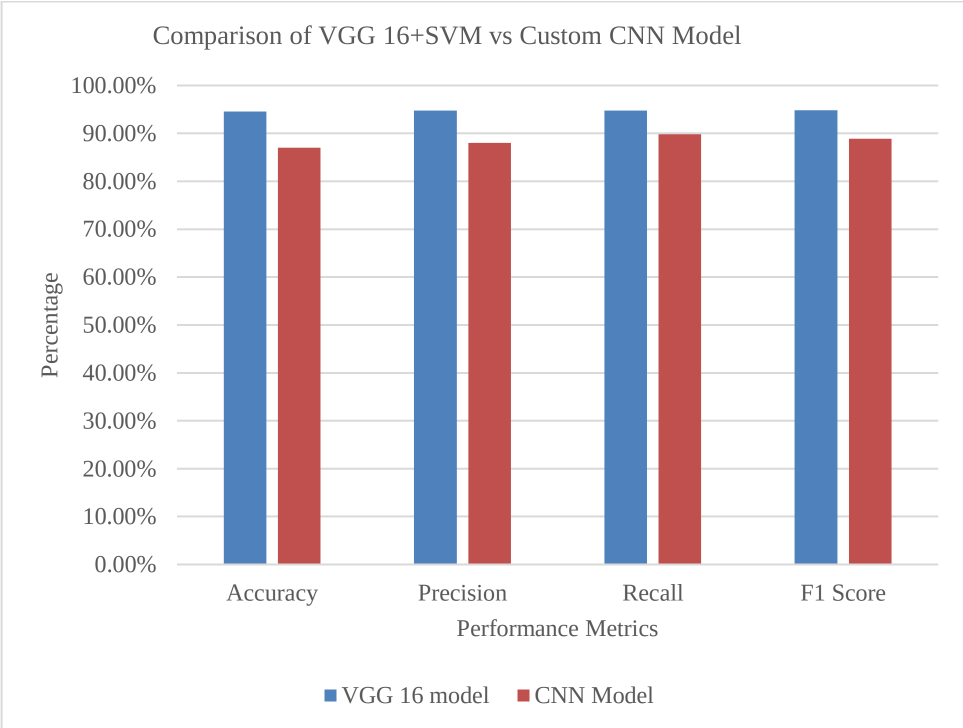

Fig 4. Comparison of VGG 16+SVM and Custom CNN Model

The performance comparison graph of VGG 16 with SVM classifier and custom CNN model is shown in the Fig.4

5. Conclusion

Our study compares the custom CNN architecture and VGG 16 architecture for detecting the presence of Pneumonia from chest X-ray Images.VGG16 achieved superior accuracy, the custom CNN presented a more efficient and flexible alternative. It is essential to investigate the interpretability and generalisability of deep learning models. For respiratory diseases like pneumonia, lung cancer, etc., a more interpretable and generalisable model is required. Our future work will concentrate on investigating better models for detecting pneumonia for providing better healthcare services.

References

1. A. Musha, Ahmmad, et al. “A Deep Learning Approach for COVID-19 and Pneumonia Detection from Chest X-Ray Images.” International Journal of Electrical and Computer Engineering (IJECE), vol. 12, no. 4, 1 Aug. 2022, p. 3655, https://doi.org/10.11591/ijece.v12i4.pp3655-3664.
2. Y. Yang, G. Mei, and F. Piccialli, "A Deep Learning Approach Considering Image Background for Pneumonia Identification Using Explainable AI (XAI)," Int. J. Electr. Comput. Eng., vol. 12, no. 4, pp. 3655-3664, Aug. 2022. DOI: 10.11591/ijece.v12i4.pp3655-3664.

3. A. M. A. Barhoom and S. S. A. Naser, "Diagnosis of Pneumonia Using Deep Learning," Int. J. Acad. Eng. Res., vol. 6, no. 2, pp. 48-68, Feb. 2022. [Online]. Available: www.ijeais.org/ijaer.
4. H. Malik, T. Anees, M. Din, and A. Naeem, "CDC_Net: Multi-Classification Convolutional Neural Network Model for Detection of COVID-19, Pneumothorax, Pneumonia, Lung Cancer, and Tuberculosis Using Chest X-Rays," Multimed. Tools Appl., vol. 82, pp. 13855–13880, 2023. DOI: 10.1007/s11042-022-13843-7.
5. N. Barakat, M. Awad, and B. A. Abu-Nabah, "A Machine Learning Approach on Chest X-Rays for Pediatric Pneumonia Detection," Digit. Health, vol. 9, 2023. DOI: 10.1177/20552076231180008.
6. D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan and A. Mittal, "Pneumonia Detection Using CNN based Feature Extraction," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-7, doi: 10.1109/ICECCT.2019.8869364.
7. Sirish Kaushik, V., Nayyar, A., Kataria, G., Jain, R. (2020). Pneumonia Detection Using Convolutional Neural Networks (CNNs). In: Singh, P., Pawłowski, W., Tanwar, S., Kumar, N., Rodrigues, J., Obaidat, M. (eds) Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019). Lecture Notes in Networks and Systems, vol 121. Springer, Singapore. https://doi.org/10.1007/978-981-15-3369-3.
8. Kermany, Daniel S., et al. “Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning.” Cell, vol. 172, no. 5, Feb. 2018, pp. 1122-1131.e9, www.cell.com/cell/fulltext/S0092-8674(18)30154-5, https://doi.org/10.1016/j.cell.2018.02.010.
9. Musha, Ahmmad, et al. “A Deep Learning Approach for COVID-19 and Pneumonia Detection from Chest X-Ray Images.” International Journal of Electrical and Computer Engineering (IJECE), vol. 12, no. 4, 1 Aug. 2022, p. 3655, https://doi.org/10.11591/ijece.v12i4.pp3655-3664.
10. Brady AP. Error and discrepancy in radiology: inevitable or avoidable? Insights Imaging. 2017 Feb;8(1):171-182. doi: 10.1007/s13244-016-0534-1. Epub 2016 Dec 7. PMID: 27928712; PMCID: PMC5265198.
11. Warren Gefter, Warren B., et al. “Commonly Missed Findings on Chest Radiographs: Causes and Consequences.” CHEST, vol. 0, no. 0, 12 Dec. 2022, journal.chestnet.org/article/S0012-3692(22)04081-8/fulltext, https://doi.org/10.1016/j.chest.2022.10.039. Accessed 2 Mar. 2023.
12. Kareem, A., Liu, H. & Sant, P. Review on Pneumonia Image Detection: A Machine Learning Approach.Hum-Cent Intell Syst 2, 31–43 (2022). https://doi.org/10.1007/s44230-022-00002-2.
13. D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan and A. Mittal, "Pneumonia Detection Using CNN based Feature Extraction," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-7, doi: 10.1109/ICECCT.2019.8869364.