iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48685

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Comparative Analysis of Deep Learning Convolutional Neural Networks based on Transfer Learning for Pneumonia Detection

Ronald Chiwariro[1], Julius B. Wosowei[2]
[1, 2]*Department of Computer Science and Engineering Jain University Bangalore, India*

Abstract: ***Artificial intelligence has been used in many different fields throughout its development, especially in recent years as the amount of data available has increased. Its major objective is to aid individuals in making more reliable decisions more quickly. Machine learning and artificial intelligence are being used more and more in medicine. This is particularly true in the medical field where a large number of digital files must be collected and processed to apply a variety of biomedical imaging and diagnostic techniques. Machine learning is used to analyse medical images, which helps with consistency and increases reporting accuracy. To process chest X-ray data and assist in diagnosis, this study compares transfer learning-based deep neural networks. The study focuses on comparing and evaluating deep learning methods based on convolutional neural networks for the identification of pneumonia. To establish the model that would accurately diagnose pneumonia, the researchers developed several different models. Five distinct CNN models, including VGG19, VGG16, ResNet50, InceptionNet v3, and YOLO v5, were trained using the RSNA Pneumonia Detection Challenge dataset. Validation Accuracy and Area Under Curve were used to gauge their performance. On test data, VGG16 had the highest validation accuracy (88%) and AUC-ROC (91.8%), whereas YOLO v5 was used to locate the inflammation with a 99% level of confidence.***
Keywords: Artificial intelligence; convolutional neural network; deep learning; transfer learning; image processing; machine learning; pneumonia detection, VGG19; VGG16; ResNet50; InceptionNet v3; YOLO v5.

I. INTRODUCTION

Pneumonia is an infection of the lungs brought on by bacteria, viruses, or fungi. It inflames the air sacs and results in pleural effusion, a condition in which the lung becomes flooded with fluid. It affects a vast population, especially in poor and underdeveloped nations where pollution, unhygienic living conditions, and overcrowding are all too common, in addition to a lack of medical infrastructure. Early diagnosis and therapy can stop the disease's progression to a fatal stage. For diagnosis, computed tomography (CT), magnetic resonance imaging (MRI), or radiography (X-rays) are frequently used to evaluate the lungs. The evaluation of the lungs using X-ray imaging is non-invasive and reasonably priced. A pneumonic can be distinguished from a medical disease by infiltrates, or white spots on the pneumonic X-ray. Contrarily, chest X-ray examinations for the diagnosis of pneumonia are subject to subjective variation. An automated pneumonia detection system is therefore required. As a result, this paper compares computer-aided diagnosis systems that use deep transfer learning models for the appropriate classification of chest X-ray pictures. Deep learning is a potent artificial intelligence method that aids in resolving numerous challenging computer vision problems. Convolutional neural networks (CNNs are frequently utilised for a variety of picture categorization problems. But when given a sizable amount of data, such models perform at their peak. Since it takes skilled medical professionals to identify each image, an expensive and time-consuming process, it is challenging to gather such a high amount of labelled data for biomedical image classification challenges. One solution to this issue is transfer learning. This method utilises models that have been trained on large datasets and applies the network weights that the models have calculated to solve problems with small datasets. CNN models are frequently used for biomedical image classification applications because they have been trained on big datasets like ImageNet [1], which contains more than 14 million images.

Using the most common technique of chest X-ray imaging and the location of the inflamed lung, the issue is to identify pneumonia. Examining chest X-rays, however, is a challenging task that is susceptible to subjectivity. In this study, we build computer-aided diagnosis methods for automatic pneumonia detection and inflammation detection from chest X-ray images. We performed deep transfer learning using VGG19 [2], VGG16 [2], Resnet50 [3], InceptionNet [4], and YOLO v5 [5].

Four traditional evaluation metrics—precision, recall, f1-score, area under the curve, and bounding box confidence levels—will be used to assess the model's performance.

©IJRASET: All Rights are Reserved

II. REVIEW OF LITERATURE

Deep learning algorithms can now analyse and segment images with accuracy comparable to that of a person. Imaging is one of the most prominent disciplines where deep learning can have a substantial impact, and it may play a key role in the medical industry. Deep learning has made enough progress to play a key role in the medical industry today. Deep learning can be applied to a variety of applications, including the brain-computer interface [9], computer-aided diagnostics, the analysis of electronic health-related data, treatment planning and drug intake, environment recognition, and the detection of tumours and lesions in medical images. The efficiency of deep learning is critically dependent on neural networks' capacity to acquire high-level abstractions from incoming raw data via a general-purpose learning algorithm [10].

While deep learning cannot currently replace doctors or clinicians in medical diagnosis, it can help medical professionals with time-consuming activities like examining chest radiographs for indicators of pneumonia. An infection of the lungs known as pneumonia is brought on by pathogens such as bacteria, viruses, and fungi [11]. Young or old, healthy or not, it can happen to everyone. Infants, people with a variety of problems, those who have a weaker immune system, the elderly, people who are hospitalised and using a ventilator, people who smoke, and people who are hospitalised and on a ventilator are all at risk. The aetiology of pneumonia affects how serious it is. Viral pneumonia has milder and more gradual symptoms.

Diagnosis of viral pneumonia with bacterial infection might be challenging. On the other hand, bacterial pneumonia is more dangerous and can have both gradual and rapid onsets of symptoms, especially in children [12]. This kind of pneumonia can spread to different lobes and damage a sizable percentage of the lungs. If multiple lung lobes are injured, hospitalisation is required [13]. Another form of pneumonia that may affect people with compromised immune systems is fungus pneumonia. The patient will need time to recover from this type of pneumonia because it can be dangerous. To reduce pneumonia-related mortality, particularly in children, in underdeveloped countries, there is an urgent need to conduct research and create cutting-edge methods of computer-aided diagnostic [14].

In medical diagnosis and therapy, the interpretation of chest radiographs is crucial. Chronic obstructive pulmonary disease (COPD) is currently the top cause of death in the United States, and by 2020, it is projected to increase [15]. The World Health Organization (WHO) reports that it is one of the leading causes of death for children under the age of five worldwide, killing an estimated 1.4 million children, or nearly 18 percent of all children under the age of five worldwide [16]. More than 90% of newly diagnosed cases of paediatric pneumonia take place in underdeveloped nations with scant access to healthcare.

Therefore, the need for low-cost and reliable pneumonia diagnostics exists. Recently, several researchers have put forth multiple artificial intelligence (AI)-based treatments for different medical conditions. Convolutional neural networks (CNNs) have been successfully applied by researchers to solve numerous medical problems, such as the diagnosis of breast cancer, the identification and segmentation of brain tumours, the classification of diseases in X-ray images, and more [17].

CNNs do incredibly well on large datasets. If proper precautions are not taken, they will commonly fail on small datasets. We conducted this study to present a novel approach to transfer learning that makes use of pretrained architectures trained on ImageNet to attain the same level of performance even on a small dataset and to detect pneumonia from routine chest X-rays. Five different pre-trained models' performance was evaluated. Finally, this study contributes by coupling a promising object detection model with the best classification model.

III. METHODOLOGY

Data was gathered from Kaggle and then put through exploratory data analysis (EDA) to look at different statistical features. The photos were pre-processed after the EDA was finished with resizing, scaling, augmentation, and annotation before being converted to jpg format. The pre-processed data was subjected to various classification and object identification models. VGG19, VGG16, Resnet50, InceptionNet, and YOLO v5 were selected for additional tweaking and final solution development.

A. Dataset

The recommended technique made use of the RSNA Pneumonia Detection Challenge [18] dataset. Which areas of the lung have symptoms of pneumonia are identified using a set of annotated chest X-rays. The stage 2 comprehensive class information, stage 2 train labels, stage 2 train images, and stage 2 test images are separated into three subfolders in the total dataset, which has 32227 images and is roughly 4 GB in size.

B. Exploratory Data Analysis

These are the specifics of the dataset files:

1) The training set may be found in stage 2 train labels.csv. It has information about the target, bounding box locations, and patientIds.

2) The file stage 2 detailed class info.csv contains specific information on the kind of positive or negative class that applies to each patientId. The class has three values depending on how the patient's lung is currently functioning: "No Lung Opacity/Not Normal," "Normal," and "Lung Opacity."

3) DICOM files (*.dcm): These are unique file types where patients' medical images are provided. They include both header metadata for the image's pixels and the underlying raw image arrays.

EDA observations shed light on the different X-rays. The X-rays shown in Figure 1 are a combination of normal, pneumonia, and non-pneumonia but not normal instances. Figures 2 and 3 show anterior/posterior (AP) and posterior/anterior (PA) X-rays, respectively.

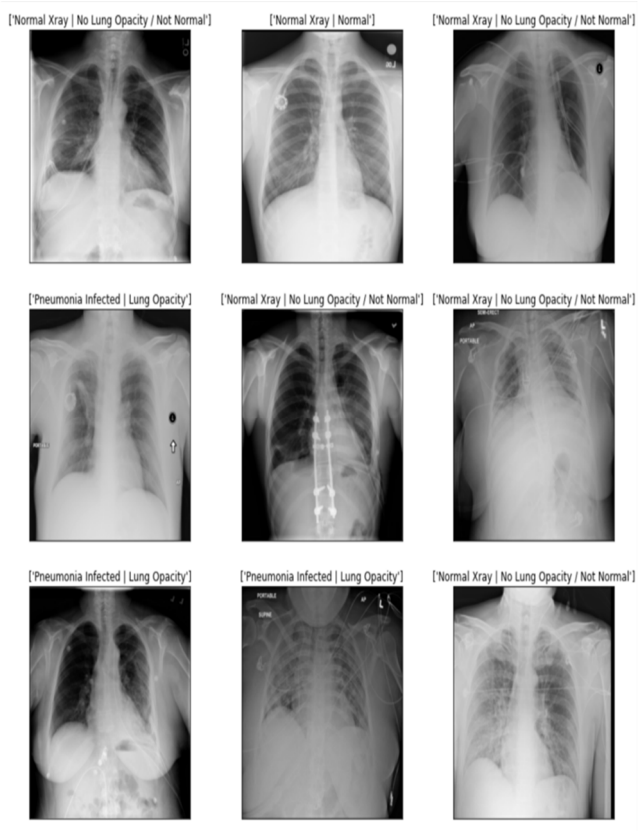


Fig. 1 Different X-ray Classes

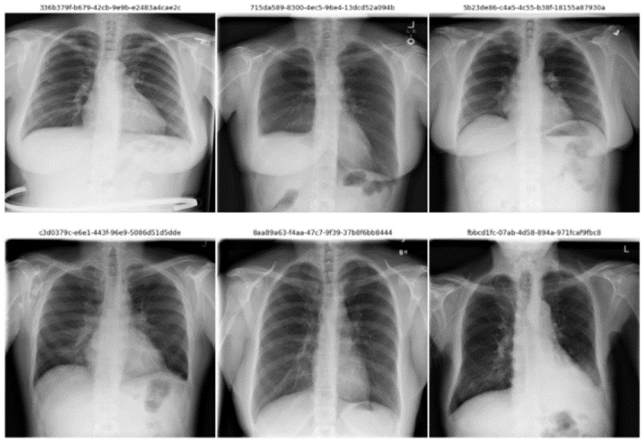


Fig. 2 Sample of Posterior/Anterior X-rays

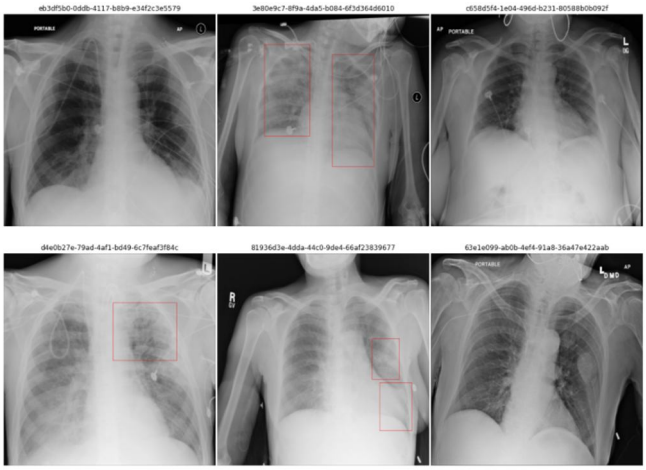


Fig. 3 Sample of Anterior/ Posterior X-rays

The sample image for a patient with no lung opacity but also not normal is shown in Figure 4. Furthermore, Figure 5 shows an image of a positive patient with Pneumonia.

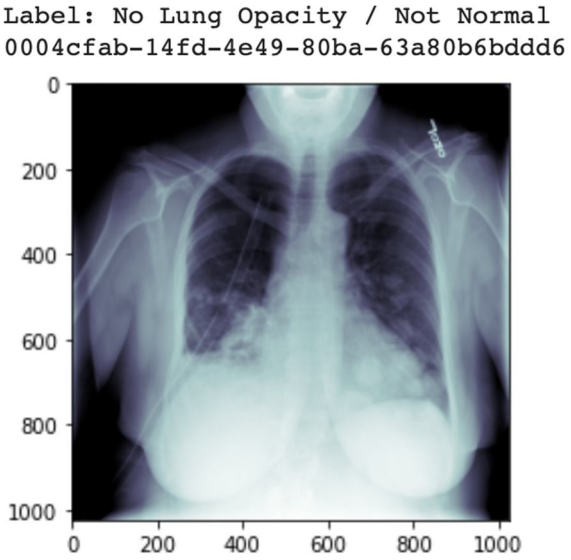


Fig. 4 Sample X-rays with no lung opacity but not normal

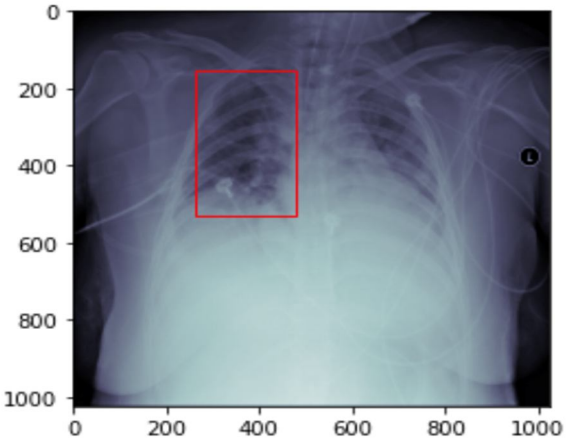


Fig. 5 Sample X-rays with Pneumonia

The insights gained from the EDA process provided a detailed understanding of the dataset in terms of the features, targets, and classes. This necessitated the pre-processing stage for the images. During this stage, the images were subjected to augmentation and annotation. This further enables the data to be passed on to the Models for training after being split into train and test sets as well as passing through the Image Data Generator.

C. Data Pre-Processing and Augmentation

To hold this dataset, each pre-trained model had to be quite huge, and each model was susceptible to overfitting. This was avoided by adding additional noise to the dataset. It is well known that sometimes significantly more generalisation can be achieved by adding noise to the neural network's inputs. Adding noise to the dataset also improves it in some way. Additionally, a variety of augmentation methods were applied. We divided the photo processing into four steps because not all augmentation approaches worked well with X-ray images.

D. Model evaluation

The process of developing a model includes model evaluation. It is beneficial to determine which model best represents our data and how well the final model will function in the future. Before choosing the optimal model, the various models were tested and assessed.

E. Classification Parameters

This multi-class classification challenge employs the categorical cross-entropy loss function. It is made to measure the variation in probability distributions between two scenarios. The accuracy of classification algorithm predictions is evaluated using a classification report. It is applied to the model's evaluation. The report displays the primary classification metrics on a per-class basis, including precision, recall, and f1-score.

The percentage of accurate predictions made using the test data is known as accuracy. It is simple to calculate by dividing the number of accurate forecasts by the total number of predictions.

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}} \quad (1)$$

Precision is defined as the percentage of instances that are truly positive and relevant among all the examples that were projected to belong in a certain class.

$$\text{Precision} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Positives})} \quad (2)$$

The percentage of examples that were correctly identified as belonging to a class in relation to all of the examples in the class is known as recall.

$$\text{Recall} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})} \quad (3)$$

The best F1 score is 1.0, while the lowest is 0.0. It is a weighted harmonic mean of precision and recall.

$$\text{F1 Score} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (4)$$

F. Object Detection Parameter

When a series of object detections from a model is compared to the actual object annotations in the dataset, Mean Average Precision (mAP) is used to assess the accuracy of the model's predictions. The mAP is computed using intersection over Union (IoU). This value, which ranges from 0 to 1, indicates how much of the predicted and actual bounding boxes overlap. IoU of 0 indicates that there is no overlap between the boxes, while IoU of 1 indicates that their union is equal to their overlap, suggesting that they are entirely overlapping. Figure 6 displays the diagrammatic representation.

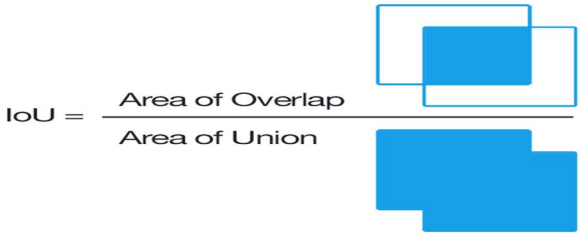


Fig. 6 Intersection over Union

IV. THE MODELS

The Models used on the dataset are VGG19, VGG16, ResNet50, InceptionNet v3 and YOLO v5. The models were trained for 50 epochs each on the same training and validation sets. Testing was done on the reserved 20% of the dataset.

A. VGG19 Model

VGG19 is a variation of the VGG model that, in essence, has 19 levels (16 convolution layers, 3 fully connected layers, 5 MaxPool layers and 1 SoftMax layer). Other VGG variations include VGG11, VGG16, and more. FLOPs in VGG19 total 19.6 billion. 224*224 RGB picture is the input for VGG-based convNet. The mean image values that are calculated for the whole ImageNet training set are subtracted from the RGB image with pixel values in the range of 0-255 in the pre-processing layer. The VGG19 Variation includes 19 weight layers, including 5 pooling layers and 16 convolutional layers with 3 fully connected layers each. Figure 7 depicts the architecture.

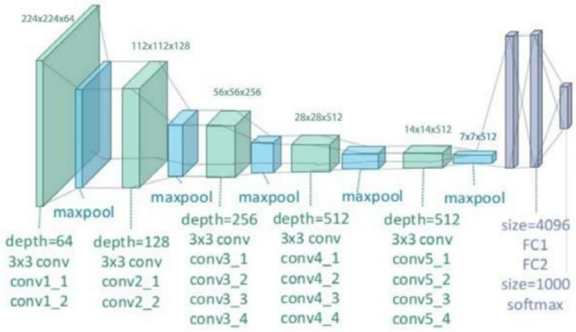


Fig. 7 VGG-19 Architecture

B. VGG16

The VGG16 CNN architecture was employed to win the 2014 ILSVR (ImageNet) competition. It is one of the best computer vision model architectures available right now. The most distinctive feature of VGG16 is that it prioritised having convolution layers of a 3*3 filter with a stride of 1 and always used the same padding and maxpool layers of a 2*2 filter with a stride of 2. Convolution and maxpool layers are arranged in this manner continuously throughout the entire architecture. The final output is provided by a SoftMax after two fully connected layers. The 16 in VGG16 stands for the 16 layers with weights. This network is quite vast and has over 138 million parameters. Figure 8 displays the architecture.

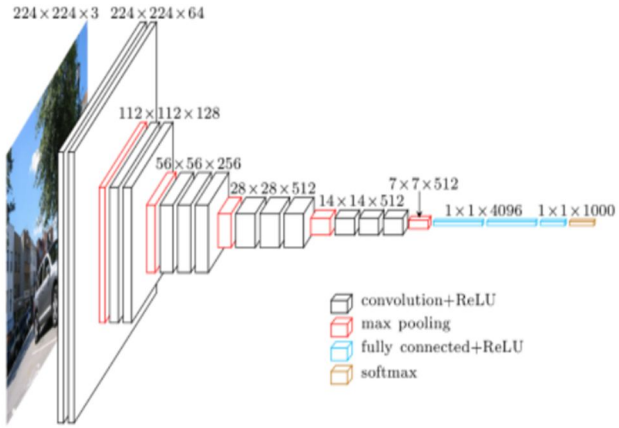


Fig. 8 VGG-16 Architecture

C. ResNet50

ResNet-50 is a convolutional neural network that is 50 layers deep. Note that there is only one 3*3 convolution rather than two. 1*1 convolutions are used to map in lower dimensions and then perform 3*3 convolution and then remap them to higher dimensions. This way the training time will be less. The other part to note is in ResNet50 when there is a dimension change then the authors used 1*1 convolutions at x to make the dimension the same, this is depicted in Figure 9.

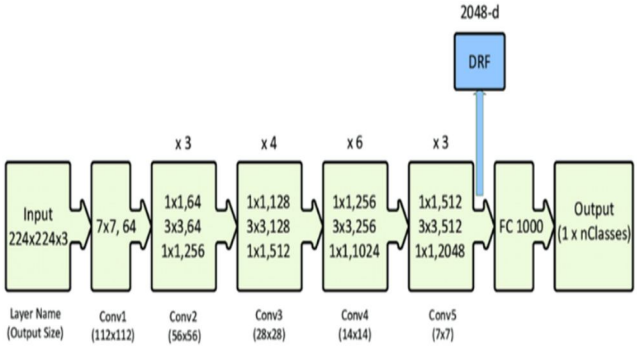


Fig. 9 ResNet50 Architecture

D. Inception v3

With the use of Label Smoothing, Factorized 7*7 convolutions, and an auxiliary classifier to transfer label information to lower down the network, Inception-v3 is a convolutional neural network design from the Inception family that offers significant advancements (along with the use of batch normalisation for layers in the side head).

Following the steps outlined below, the architecture of an Inception v3 network is gradually built:

1) Factorized convolutions: By lowering the number of parameters used in a network, this method aids in decreasing computing efficiency. The effectiveness of the network is also monitored.

2) Smaller convolutions: Training is completed more quickly by substituting smaller convolutions for larger ones. A 5*5 filter, for example, contains 25 parameters; two 3*3 filters, in contrast, have only 18 (3*3 + 3*3). Figure 10 illustrates architecture.

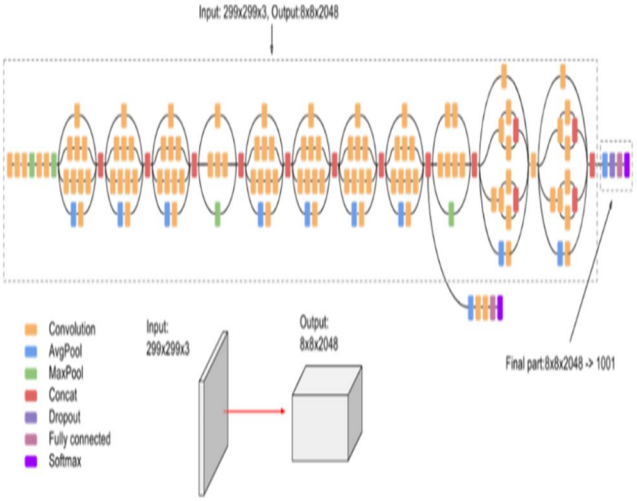


Fig. 10 Inception v3 Architecture

E. YOLO v5

YOLO uses a single neural network to predict bounding boxes and class probabilities from complete photos in a single evaluation. The whole detection process is a single network, allowing for end-to-end optimization-based solely on detection performance.

The batch inputs are m photos of shape (m, 416, 416, 3). A convolutional neural network receives this image from the YOLO model. The output volume is (19, 19, 425) after flattening the last two dimensions.

In this case, a 19*19 grid yields 425 results in each cell.

1) 5 is the number of anchor boxes in a grid, hence 425 is calculated as 5*85.

2) 85 = 5+80, where 5 represents (pc, bx, by, bh, and bw) and 80 represents the number of classes we want to identify.

The output consists of a list of bounding boxes and the identified classes. Six digits represent each bounding box (pc, bx, by, bh, bw, c). Each bounding box has 85 values to represent it if we expand c into an 80-dimensional vector.

As illustrated in Figure 11, we perform the IoU (Intersection over Union) and Non-Max Suppression as our final step to prevent picking overlapping boxes.

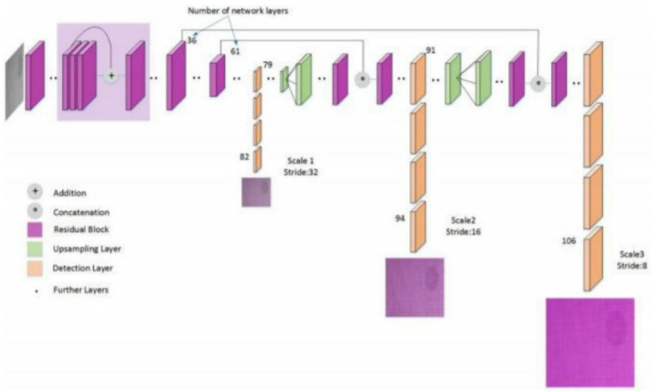


Fig. 11 YOLO v5 Architecture

V. RESULTS

The four Models namely, VGG19, VGG16, ResNet50, InceptionNet v3 and YOLO v5 were trained on the full dataset and their performances based on Validation Accuracy and Area Under Curve (AUC) were as follows:

1) *VGG19:* Validation Accuracy - 80% and AUC-ROC – 82.5%.
2) *VGG16:* Validation Accuracy - 88% and AUC-ROC – 91.8%.
3) *ResNet50:* Validation Accuracy - 73% and AUC-ROC – 80.9%.
4) *InceptionNet v3:* Validation Accuracy - 79% and AUC-ROC – 87%.
5) *YOLO v5:* Inflammation-detection bounding boxes with high confidence.

Therefore, it can be observed that the VGG16 Model has the highest Validation Accuracy and Area Under Curve for Classification. YOLO v5 is then applied to locate the position of the inflammation.

Comparison of Classification Reports

VGG19

	precision	recall	f1-score	support
0.0	0.95	0.81	0.87	107
1.0	0.44	0.76	0.56	21
accuracy			0.80	128
macro avg	0.70	0.79	0.72	128
weighted avg	0.86	0.80	0.82	128

VGG16

	precision	recall	f1-score	support
0.0	0.97	0.88	0.92	107
1.0	0.58	0.86	0.69	21
accuracy			0.88	128
macro avg	0.77	0.87	0.81	128
weighted avg	0.91	0.88	0.88	128

ResNet50

	precision	recall	f1-score	support
0.0	0.86	0.86	0.86	102
1.0	0.46	0.46	0.46	26
accuracy			0.78	128
macro avg	0.66	0.66	0.66	128
weighted avg	0.78	0.78	0.78	128

Inception v3

	precision	recall	f1-score	support
0.0	0.84	0.78	0.81	90
1.0	0.56	0.66	0.60	38
accuracy			0.74	128
macro avg	0.70	0.72	0.71	128
weighted avg	0.76	0.74	0.75	128

Comparison of *Area Under Curve*

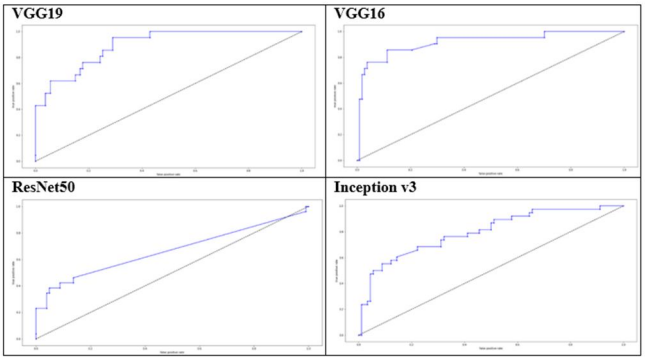


Fig. 12 Classification and Area Under Curve Results

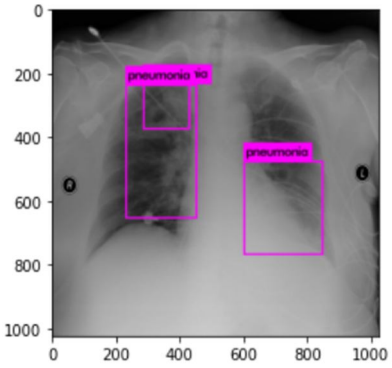
A. *Location of Inflammation*

After an image is classified to be pneumonia positive, it undergoes object detection with YOLO v5 to locate the position of inflammation. Figures 13 and 14 show the results of positive images subjected to YOLO v5 after being classified positive for pneumonia.

```
#Plot the prediction with the position of Lung Inflammation

plt.imshow(cv2.imread("predictions.jpg"))
```

```
<matplotlib.image.AxesImage at 0x7fba6798d250>
```



- Detections of inflammations on test images.
- Default Threshold of 0.005% was used during testing.

Fig. 13 Detection of inflammation in pneumonia-positive patients

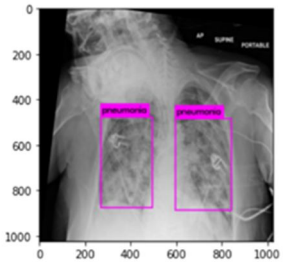
```
./images/395738ff-53d2-42e6-aa6d-a01e4a07f01f.jpg: Predicted in 22.544326 seconds.
pneumonia: 99%
pneumonia: 96%
```

- Boxes show confidence of:
  - 99%
  - 96%

```
#Plot the prediction with the position of Lung Inflammation

plt.imshow(cv2.imread("predictions.jpg"))
```

```
<matplotlib.image.AxesImage at 0x7fba67a881d0>
```



- Detections of inflammations on test images were correctly done with a high confidence level.
- Threshold of 25% was applied as a filter during testing.

Fig. 14 Detection of inflammation in pneumonia-positive patients

VI. CONCLUSION

In this study, deep learning was used to classify digital chest X-ray images according to whether or not they showed changes that were consistent with pneumonia. Python was used to carry out the implementation, which was based on the CNN model. For the correct diagnosis of any ailment, the input and presence of medical specialists are still necessary. To build a reliable and robust illness categorization model, it is essential to gather as much data as possible. The project's next steps include experimenting with different CNN and pre-processing configurations, data augmentation techniques, and using additional X-ray datasets with supplementary data labels showing other pathologies. In the classification of pneumonia and non-pneumonia patients, VGG16 performed better than other classifier-based Models, and YOLO was used to identify objects to determine the precise site of the inflammation. Pneumonia cases tend to be more prevalent in men over the age of 40, and all cases were correctly identified with high confidence levels in the bounding boxes.

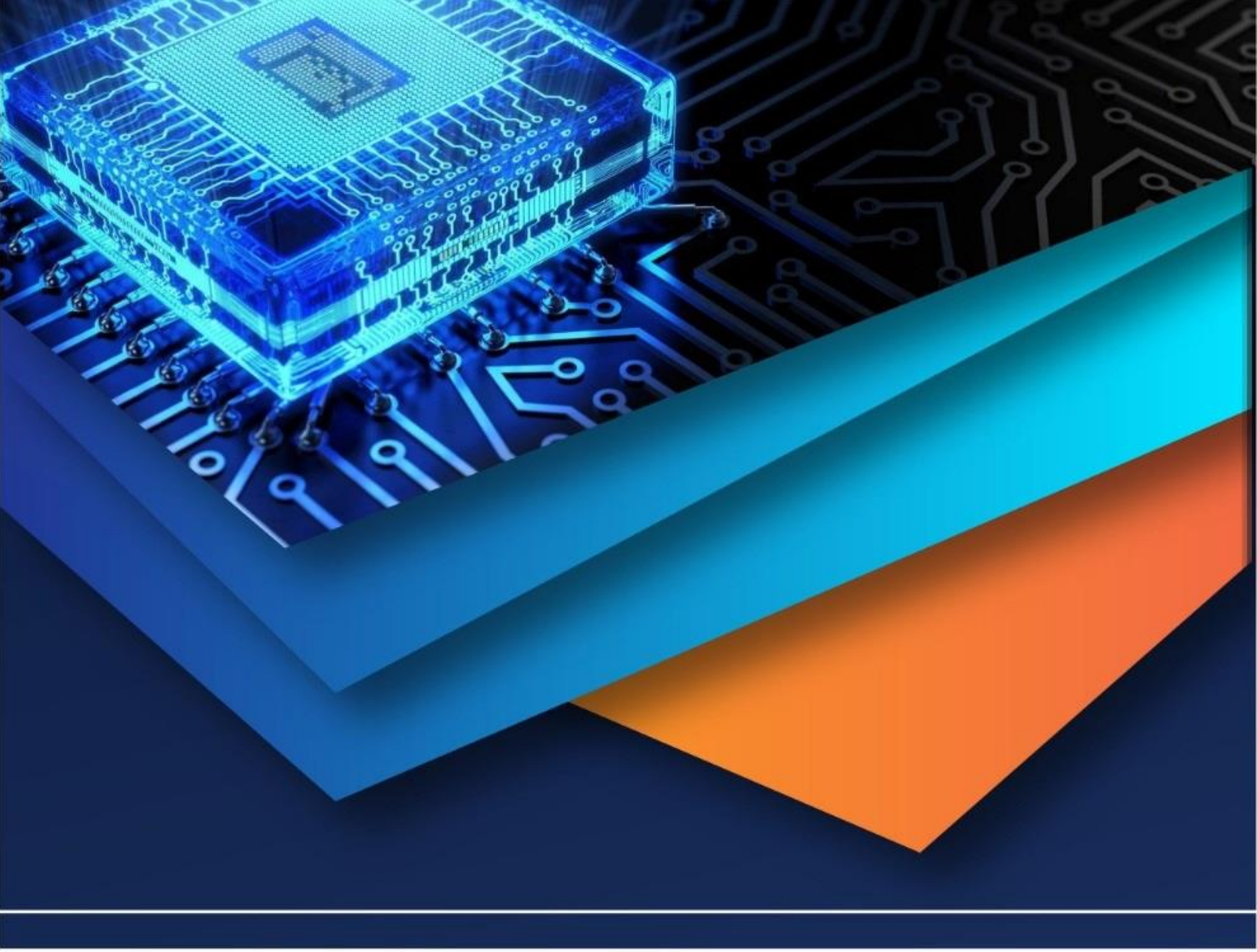
It is concluded as a result that a combination of VGG16 and YOLO v5 can be used successfully by medical professionals for the early detection of pneumonia in both children and adults. Healthcare organisations can provide more effective patient care and reduce mortality rates by quickly analysing a large number of X-ray images to obtain incredibly precise diagnostic results. These convolutional neural network models were built using a variety of parameter tuning techniques, such as dropout, altered learning rates, altered batch size, altered number of epochs, added more complex fully connected layers, and altered stochastic gradient optimizers.

VII. FUTURE WORK

The backbone network of the existing detection algorithms still has issues. For instance, Resnet often has two issues: significant down sampling that results in the loss of the target position and semantic information and a huge network depth that prolongs training time.
The dataset has three categories: normal, pneumonia, and abnormal (cancer or other disorders), but only the bounding boxes for photos of pneumonia are provided. Resnet50 was unable to differentiate between pneumonia and abnormal images because the characteristics of abnormal (cancer or other disorders) and pneumonia are identical. As a result, it is possible to forecast the bounding box for aberrant photos.
Generally speaking, it's not essential to specifically mention "negative pictures". TensorFlow does not perform well during training in this instance since a significant chunk of the data set contains bad examples. Negative picture detections are not penalised by the loss function. As a result, the model frequently indicates a case of pneumonia even though the lung is not normal.

REFERENCES

[1] Liu, N.; Wan, L.; Zhang, Y.; Zhou, T.; Huo, H.; Fang, T. Exploiting Convolutional Neural Networks With Deeply Local Description for Remote Sensing Image Classification. IEEE Access 2018, 6, 11215–11228.
[2] https://keras.io/api/applications/vgg/.
[3] https://www.kaggle.com/keras/resnet50.
[4] https://keras.io/api/applications/inceptionv3/.
[5] ultralytics/yolov5: YOLOv5 in PyTorch , ONNX, CoreML, TFLite (github.com).
[6] Shickel, B.; Tighe, P.J.; Bihorac, A.; Rashidi, P. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE J. Biomed. Health Inform. 2018, 22, 1589–1604.
[7] Meyer, P.; Noblet, V.; Mazzara, C.; Lallement, A. Survey on deep learning for radiotherapy. Comput. Biol.
[8] Malukas, U.; Maskeliunas, R.; Damasevicius, R.; Wozniak, M. Real-time path finding for assisted living using Deep Learning.
[9] Zhang, X.; Yao, L.; Wang, X.; Monaghan, J.; McAlpine, D. A Survey on Deep Learning based Brain Computer Interface.
[10] Bakator, M.; Radosav, D. Deep Learning and Medical Diagnosis: A Review of Literature. Multimodal Technol. Interact. 2018, 2, 47.
[11] Gilani, Z.; Kwong, Y.D.; Levine, O.S.; Deloria-Knoll, M.; Scott, J.A.G.; O'Brien, K.L.; Feikin, D.R. A literature review and survey of childhood pneumonia aetiology studies: 2000–2010. Clin. Infect. Dis. 2012, 54 (Suppl. 2), S102–S108.
[12] Bouch, C.; Williams, G. Recently published papers: Pneumonia, hypothermia and the elderly. Crit. Care 2006, 10, 167.
[13] Scott, J.A.; Brooks, W.A.; Peiris, J.S.; Holtzman, D.; Mulholland, E.K. Pneumonia research to reduce childhood mortality in the developing world. J. Clin. Investig. 2008, 118, 1291–1300.
[14] Wunderink, R.G.; Waterer, G. Advances in the causes and management of community-acquired pneumonia in adults. BMJ 2017, 358, j2471.
[15] Heron, M. Deaths: Leading causes for 2010. Natl. Vital. Stat. Rep. 2013, 62, 1–96.
[16] World Health Organization. The Top 10 Causes of Death; World Health Organization: Geneva, Switzerland, 2017; Available online: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed on 10 November 2019).
[17] Kallianos, K.; Mongan, J.; Antani, S.; Henry, T.; Taylor, A.; Abuya, J.; Kohli, M. How far have we come? Artificial intelligence for chest radiograph interpretation. Clin. Radiol. 2019, 74, 338–345.
[18] https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/overview



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)