a

International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

A Comparative Study on Pre-Training Models of Deep Learning to Detect Lung Cancer

Madhavi Aluka[*1], Sumathi Ganesan[#2], Vijay Pal Reddy P[3]

Submitted: 28/10/2022 **Revised:** 10/12/2022 **Accepted:** 31/12/2022

Abstract. Detection of lung cancer using neural network-based systems has seen a reasonable improvement. However, the possibility of false cancer detection seems to be a worrying factor in recent times due to various technical reasons. Recent research programs revealed that machine learning (ML) based techniques were also found to make a greater contribution to lung cancer detection. However, deep learning (DL) techniques seem to provide enhanced accuracy for various medical research areas. Therefore, in this work different types of DL pre-trained prediction models are tested to study the accuracy of each model. The pre-trained models are applied to the dataset consisting of nearly 5000 images consisting of cancerous and non-cancerous data. Particularly, VGG-16, Inception V3, and ResNet50 are the Transfer Learning models used in this study for comparative analysis. The results show a reasonable accuracy using the VGG-16 model with fine-tuning and the image augmentation obtained greater accuracy of 96% and 93% for training data and validation data respectively.

Keywords: *Deep Learning, VGG-16, Inception V3, ResNet50, Lung Cancer, Pre-trained Models.*

1. Introduction

Transfer learning is frequently expressed using pre-trained models. It has recently captivated academics' and scholars' curiosity, and it has been well-implemented in various domains. ImageNet data, a vast collection of image data encompassing 1000 types of images, was utilized to train pre-trained image categorization. These models have been pre-trained to classify any image that falls into one of the 1000 categories.

This method involves transferring a portion of the network on trained model to comparable job and then retraining the model to classify by adding one or more layers at the end. These are saved Deep Neural Network models with their weights after training, images are trained on an exceptionally large dataset for classification, and the lower layers of convolutional neural networks (CNN) learn the fundamental characteristics of the dataset.

CNNs are composed of layers, the input layer, the convolution layer, the pooling layer, the fully connected layer, softmax layer and the output layer. CNN is designed into two sections, 1. For analysis, convolution techniques extract and define the many aspects of a lung image, 2. Using data extraction from earlier phases, Convolution is performed to predict the outcomes of the classes in this layer. Use a lung image as input with pixel intensity, apply filter to an original image and initialized with random weights. These weights are to be learned by network (CNN) and place the filter on the top left of the input image and perform convolution. multiply the pixel intensities with the corresponding value on the filter and then take the max value and update the central pixel in the region where the filter is placed. After performing convolution only central pixel is updated and others will be zero. Move the filter from left to right and when it reaches the rightmost pixel, move it one pixel down, from left to right again, and perform this on the whole image. the resultant of the operation, convolution image is less than the input image.

After extracting the features from the lung image, the full connected layer is used for the classification purpose. This layer classifies the lung cancer or non-lung cancer images based on the feature extraction. Then finally the dense is the output layer which displays whether the lung image is cancerous or non-cancerous.

It is the most generalized topic and is considered as a subset of ML, DL, and artificial intelligence (AI). A mimic of the human brain is created by the network layers and based on the pixel value it tries to identify the features of lung images. As a result, these bottom layer characteristics are almost identical for classification. As a

[1]*Department of Computer Science Engineering, Annamalai University, Chidambaram, Tamil Nādu, INDIA*
ORCID ID : 0000-0002-8440-025X
[2] *Annamaliai University, Chidambaram, Tamilnadu, INDIA*
ORCID ID : 0000-0001-6696-4003
[3] *Matrusri Engineering College, Hyderabad, Telangana, INDIA*
ORCID ID : 0000-0002-0534-2187
[*]*Corresponding Author Email: alukamadhavi2000@gmail.com*

* *Corresponding Author Email: author@email.com*

result, the pre-trained model's initial layers learn the characteristics for identifying lung cancer. It is built on a generalized dataset. A deep Neural network's accuracy increases with the size of the dataset When the dataset is insufficient, to get accurate results, then use a pre-trained model which already learned the relevant characteristics from the Architecture. Pre-trained models are as listed here i) VGG-16, ii) Inception V3 and iii) ResNet 50 were studied and examined for the enhancement of accuracy of lung detection.

1.1. Dataset

Dataset is a part of the lung Image Database Consortium and the Infectious Disease Research Institute (LIDC/IDRI). LUNA16 can be used with a reformatted version of these images since they were very large (124 GB). There were 888 CT images in the dataset, each with annotations providing coordinates and ground truth labels, refers the accuracy which is used by the training set for classification. In statistical models, this is used to confirm or invalidate research hypotheses.

The 64 channels with 3X3 filter size of the first two layers of a VGG 16, just create convolution layers with 128 filter size. The max pool layer stride (2,2) (3,3) is identical to that of the stride with (2,2) max-pooling layer. Many residual networks help in solving the degradation problems by skipping connections and short-circuiting shallow layers in the deep layers. With the backing of residual blocks, there will not be any issue of performance degradation and enables to function with high stability. Whereas in case of inception module, the image models aim to obtain the approximate optimal local sparse in CNN. This simply allow multiple types of filter sizes and allows the information to pass on to next layers.

1.2. Create an Image Database

By using this procedure, the DICOM file has been converted into a raw file (multidimensional image data) and .mhd file. The image information is extracted from the raw file following the file format conversion, and the non-image information is placed in the mhd header file. Each of these CT scan images has size of $512 \times 512 \times n$, here n is the number of axial scanned images. Approximately 200 images are present in each CT scan. There were 551065 annotations. 1351 of the annotations were labelled as nodules, while the rest of them were tagged as negative. As a result, there will be a wide class imbalance. The simple solution is to sample the majority class and augment the minority with images that are in rotated format. The pre-trained models might be trained on all pixels, but this would result escalated in cost-related issues due to delayed training time. Instead, simply crop the images around the annotations and Cartesian coordinates.

As a result, they needed to be converted to voxel coordinates and the image intensity had to be specified in Hounsfield scale. As a result, it was rescaled for image processing. The script would generate grayscale images of 50×50 for testing, training, and validation sets a pre-trained model. There were images of class-like (growth of Abnormal tissue) in one out of every six images. For training, the data set is still wildly unbalanced. then, add rotating images to the training set. The augmentation resulted in a class distribution of 80-20, which was not optimum. However, it could result in a minority class with some differences. So, it is overfitting, which is a technique used to solve an imbalanced dataset to augment the minority class. Each CT scan has dimensions of 512 × 512 × n. Cropping the images specified in the annotations would increase the computational cost and training time, as would training the pre-trained models, These annotations were provided in Cartesian coordinates, which were converted to coordinates, and the intensity of the images was defined in Hounsfield scale, The Hounsfield unit (HU) is a relative quantitative radio density measurement used by radiologists to interpret computed tomography (CT) images. which required rescaling for image processing. For training, testing, and validating a pre-trained model, the script would generate 50 × 50 grayscale images. It also splits the dataset into a training set (80%) and a test set (20%).

1.3. Leveraging the Pretrained Model with CNN Models

It is usually beneficial to reuse or build new models from the pretrained models for the following two reasons. 1. Feature extraction and 2. Finetuning.

1.3.1. Feature Extraction

As part of the size depletion process, raw data is chopped up into smaller chunks for processing. Most of these variables in these large data sets require a lot of computational resources to process.

1.3.2. Fine Tuning

This is a way of applying or utilizing transfer learning. Except for the output layer, the model transfers all designs and the model tunes the parameters based on the information it receives from the source.

2. VGG-16

VGG-16 is called pre-trained as it is already trained on large datasets like ImageNet. Which is the collection of various categories with such images. Using this, the model should have learned one of the robust hierarchies of the features that are spatially, rotationally, and translationally invariant when compared to a CNN model. The model is also known to work better for feature extraction, which includes new images. This is also an advantage to be used for Computer Vision purposes. ImageNet may have categories of images; the new images may not fall in those categories. But the features in them can be extracted well when used better.

2.1. VGG-16 Architecture

VGG-16 Model Architecture includes 13 convolution layers, and it also uses 3×3 In addition, it uses convolution filters with maxpooling layers for downsampling [1]. These two are fully connected to the hidden layers and each layer is 4,096 units. In the preprocessing layer, pixel values r aging from 0 to 255 are captured from the RGB image [2]. Then subtracts the image value which is calculated during the training set. The dense layer consists of 1,000 components, in this each unit is performed as one of the image categories in the ImageNet database [3]. This layer is the input to the convolution layer passing the imge set for training through it.

Convolution layers with 64 filters make up the first two layers. The overall volume will reduce by using the pooling layer.

Using fully connected dense lyers for the image prediction eliminates the need for the last 3 layers. The first five blocks are very important as they are involved in an effective way of feature extraction [4]. In the last model, the fine-tuning for the VGG model includes the automatic weights and the data used for training. VGGNet has 19 weight layers of which 16 are the convolution layers including 3 layers are fully interconnected.

Similarly, there are 5 pooling layers with 1000 channels including 100 labels. In VGGNet there are 2 layers are feully interconnected, and it contains 4096 channels each which are succeeded by another fully connected layer. Layers are fully interconnected uses the softmax layer which helps the model for classifying functions [5].

VGG-16 is widely used in architecture as it has 16 layers, including a convolution base layer that includes a sequence of conv2D max pool and conv2d pooling. After that, it uses a flattened layer for transforming the matrix into a vector and adds dense layers for classification. As a result, the upper dense layer is referred to as a classifier. which consists of a convolutional base and an actual classifier. It uses the features for classification and is loaded with the image net dataset with pre-trained weights. Imagesnet contains thousands of classes, of which two are discussed in this paper. The upper layers should be modified so that the output layer contains only one neuron. if the input image is cancerous or not cancerous, and trains the modified neural network from beginning to end. But keep in mind that this convolution base must be frozen and use the features for the model. Fine-tuning and augmentation strategies are being used on the lung images in VGG-16 for increasing the accuracy and to reduce the losses as VGG-16 tends to learn from images of lung images bitwise.

2.2. Understanding the VGG-16 Model

This VGG-16 model architecture as shown in Fig.1, has around 16 layers that are fully connected [6]. The model is trained on ImageNet which is used for both image classification as well as Computer Vision.

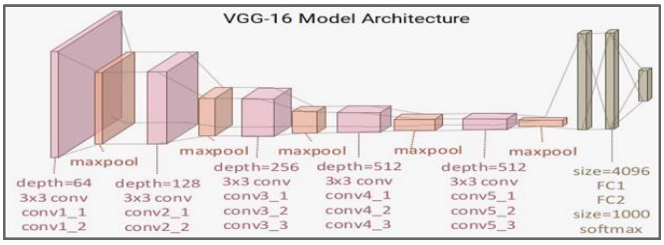

Fig. 1. VGG-16 Model Architecture.

The VGG-16 model's architecture includes 13 convolution layers with 3×3 convolution filters, as well as max pooling layers for downsampling, two fully connected and hidden layers with 4096 units each, and a dense layer with 1000 units. Where down sampling is a disproportionately small sample of the majority class for training, ImageNet’s image categories are shown in each unit. In order to forecast whether images are malignant or not, already use fully connted dense layers.Our model focuses on the first five blocks, leveraging As a functional feature extractor, the vgg model can be used. One of the 5 convolution blocks uses a feature extractor by freezing the blocks after every epoch to ensure the weights are not updated. After this step.VGG model fine tuning to, Blocks 4 and 5of the VGG model are unfrozen to allow the updating process of each epoch at the time of training the model. The model training includes two parts, the feature extractor which is basic, and the tuning part.

2.3. Feature Extraction with Pretrained CNN Model

Model VGG-16 is loaded using the Keras and teg convolution block of layers seen in Fig.2, this way the model is used as a feature extractor. To extract features, the network consists of an input layer, three convolutional and average pooling layers, and a softmax fully connected output layer. For classification, a two-layer hidden neural network is utilized once features are extracted. Three fully connected layers at the apex are included in VGG16 (top function), weights are predefined weights, 'imagenet is random initialization' (pre-training on ImageNet). The image's dimension is given by input shape.

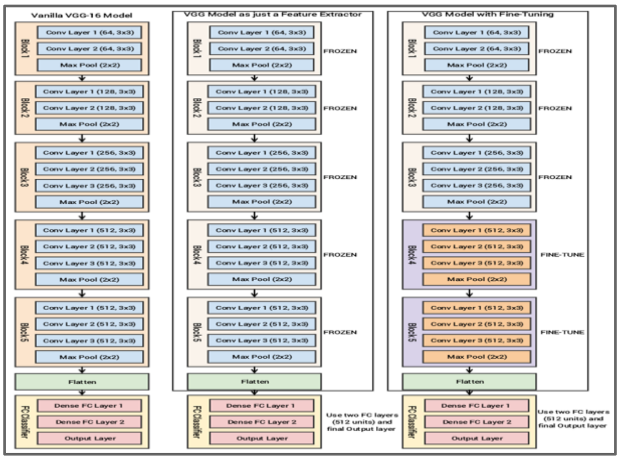

Fig. 2. VGG-16 Model Description.

The previous result shows that, the VGG-16 model's layers are frozen, which is a good thing, and it does not require the model's weight at the time of training the models. The bottleneck features come from the VGG-16 model's last activation feature map (output from block5 pool), which may then be flattened and used to signify the fully connected classifier based on a deep neural network, Converting the size of the input image into the output image size.The bottleneck characteristics for a sample image to train data.

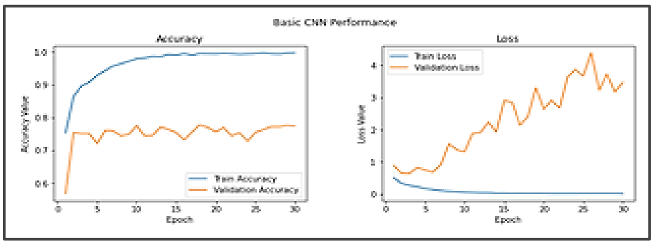

Fig. 3. VGG-16 Basic Model.

Flatten the bottleneck features of the VGG model object so that to fed the fully interconnected classifier. The model's advantage is that the training set's deatures can be extracted and then used as input for the current classifier. So that the validation sets are also affected by this, consider the training and validation sets that are currently available, then identify the bottle neck features. This model could successfully extract the flattened bottleneck features 1×8192 dimensions for the given training images of 5,000 and 1,000 validation images as shown in Fig. 3 For the deep neural network based classifier, the architecture is designed.To which the features can be considered as an input.

Table 1. Training and Validation of VGG-16 Model.

Training Data		**Validation Data**	
Accuracy	Loss	Accuracy	Loss
92	0.2	88	6

The validation accuracy is of 88 %, shown in Table 1. which is nearly a 5–6% upgraded version over the CNN model which is basic and with image augmentation. [2], which is exceptional. However, the model appears to be overfitting. After the fifth epoch, there is a significant gap between validation accuracy with model train, which indicates overfitting of training data. However, with regard to accuracy, this model appears to be the best. Lets use this model to test image augmentation methodology..First, save model to disc with the code (model. Save) with .h5 Extension Hierarchical Data Format (HDF) An H5 file is an HDF-formatted data file. It contains scientific data in multidimensional arrays and is used in Engineering, academic research, astronomy and the medical field.

2.4. Combination of Image Augmentation and Pre-trained CNN Model for the Feature Extraction

To Train and validation datasets make use of data generators that are comparable to inherit from the Sequence class to create a custom data generator to accomplish that. Adding the arguments to the Sequence class forces us to create two methods: __len__ and __getitem__, additionally to the role towards the end of the epoch. For understanding, the code used to create the model is shown below after each epoch, therefore, having bottleneck features like in the previous model is not absolutely necessary.Whereas in the VGG16 model, bottleneck features are the last activation maps before the fully connected layers, given that it is train with data generators, which will pass the VGG-model object as input to the proposed model. As a result of the model having to train for 100 epochs in this place .no sudden weight adjustments are required to the proposed model layers because the learning rate has slightly gone down. Here, VGG-16 model is used as feature extractor.

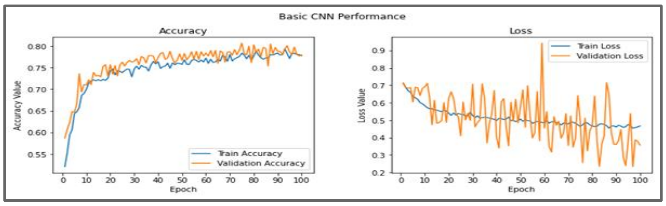

Fig. 4. VGG-16 Model with Augmentation.

Table 2. Training and Validation of VGG-16 with Augmentation.

Training		**Validation**	
Accuracy	Loss	Accuracy	Loss
78	0.5	90	6

Compared to the prior model, this has a greater validation accuracy of 90%as shown in Table: 2. It is also noted that to train and validation accuracy of this model are quite near to one another, which indicates no over-fitting. Save this model to hard disc to test and compare the result. Block 4 and 5 in the previous block diagram should be unfrozen in order to optimise the VGG16 model and build the classifier.

2.5. Application of Fine-Tuning and Image Augmentation on the CNN model

Fine-tuning and image augmentation is used to train the neural networks from beginning to end, unfreeze some of the hidden layers of frozen parts, and modify the higher actual classifier neural network. As a result, in the higher layer of convolution base and the actual classifier part, some parts will be frozen, and others will be unfrozen to learn abstract and complex higher layer features for the task for classifying cancer and non-cancer. To consider the better sampling approach and increase the size of the training set, image augmentation is useful. Convolutional neural networks will perform better by increasing the number of lung images in the training set.Using Image DataGenerator, the image in a way that retains the important features for making predictions.

Leveraging the VGG-16 model, fine-tuning a network is altering the parameters of a previously trained network so that it adjusts to the new task. As previously said, the layers which are at the start are to learn very basic features, as the model is moved further, the pattern in the images becomes clear. This helps in training the layers with the patterns specifically. Block 4 and 5, are used here for training, whereas the other 3 models in the front are kept frozen.

The convolution and pooling layers training specifications for Blocks 4 and 5 in see Fig. 2, as seen by the above output. This describes the effect of backpropagation which updates the layer weights after each epoch, on the model. By employing the same data generators and model architecture to train the model as in the previous. Since a sudden update of the weights is not preferred for the trainable VGG16 model and could negatively affect the model ,the code wont get stuck at any local miniums

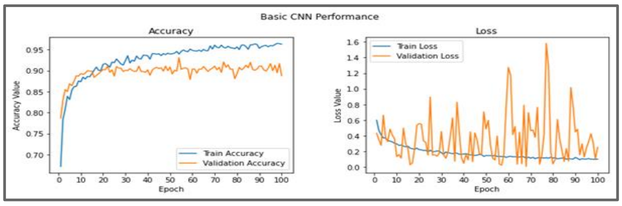

Fig. 5. Accuracy and Loss of the Pre-trained Model when Fine-Tuning and Image Augmentation is applied.

Table 3. Training and Validation of Pre-trained Model with Fine-Tuning and Image Augmentation.

Training		**Validation**	
Accuracy	Loss	Accuracy	Loss
96	0.2	93	6

According to table: 3, the current model's test accuracy is approximately 93%, which represents a 6% improvement over previous model. Overall, this model's test accuracy has increased by 24% when compared with basic CNN Model. It an example of how transfer learning is valued and accuracy values can be found and even though the model appears to be moderately overfitting on the train data, it has acceptable test accuracy. Now use the following code to save this model to disc by model. Save ('lung_cnn_tlearn_finetune_img_aug_cnn.h5').

3. Evaluations of the Model on The Test Data

It certainly yields some intriguing results. Each subsequent model showing some of the better results than previous model. This is a result of cutting-edge methods used in model. One of the instances of a bad fit in comparison to the best model with finetuned model that has transfer learning and image augmentation, cnn model has accuracy and f1score of about 72%. The accuracy and F1 score of this model were 93%. In addition, this model has been taken into account for training the model using a train dataset of 5000 images. The summary of model classification is listed in Table 4 with and without fine tuning.

Table 4. Model classification of lung images with and without fine tuning.

Model Classification Report for Lung Image Augmented and Feature Extractor

	Precision	**Recall**	**F1-score**	**support**
Cancer	0.56	0.69	0.62	225
Not cancer	0.89	0.83	0.86	700
Accuracy			0.79	925
Macro average	0.73	0.76	0.74	925
Weighted avg	0.81	0.79	0.80	925

Model Classification Report for Lung Image Augmented and Fine Tuning

	Precision	Recall	F1-score	support
Cancer	0.90	0.67	0.77	225
Not cancer	0.90	0.98	0.94	700

Accuracy			0.90	925
Macro Average	0.90	0.82	0.85	925
Weighted Avg	0.90	0.90	0.90	925

The following section compares the ROC curves of each of these models. A receiver operating characteristic curve is a graph that shows the sensitivity and specificity of a classification model for a given set of tests. On this graph, two parameters are displayed: True Positive Rate and False Positive Rate. This demonstrates the distinction between transfer learning and pretrained models, particularly when handling complex problems with limited data. shown in Fig. 6.

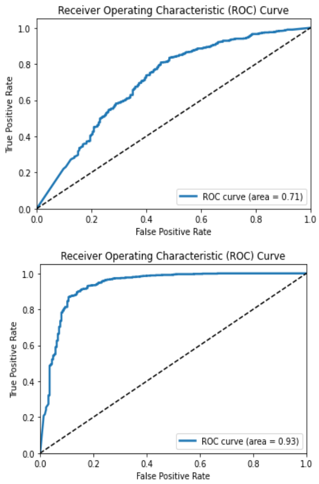

Fig. 6. Difference between Normal Model and Transfer Learning.

4. Inception V3

Inception V3 is one of the CNN architectures that belongs to the inception family. The structure of these layers is different when compared to other models. It helps in reducing the overfitting and computation power of the training data. Convolution of 1 × 1 is used with 128 filters for the purpose of dimension reduction and rectified linear activation [7]. An auxiliary classifier is used as regularization in the model [8]. In the architecture of Inception V3, Lung Images that already been used with first layer, which is the previous layer. On the image, different types of convolutions and different types of filters are applied. Firstly, the 1×1 filter is applied to the image and convolution is performed after the process is completed then feature maps are formed. Similarly, the next filters 3×3 is applied on the image and the convolution process is continued after that feature maps are formed. Again, the 5x5 filters are applied on the image and convolution is performed. After the process is completed, once again the feature maps are formed. After completion of applying the three filters then another 3×3 filters are also applied on the image and pooling process is performed on it and later feature maps are formed. Many different filters are applied to form a feature map. The feature map and the lung image will both be the same size. The process is repeated 3 times, 4 times and 2 times Based on the model. The same building blocks are processed in subsequent steps and the design of the inception model is the result of this. Inception V3 is widely used as an image recognition model and has shown 78.1% accuracy. This, when compared to another model, is good.

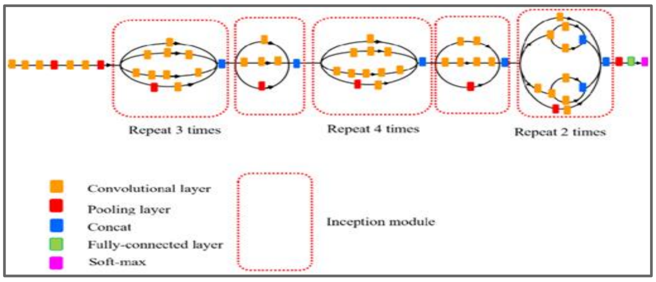

Fig. 7. Inception V3.

Table 5. Training and Validation of Inception V3.

Training		Validation	
Accuracy	Loss	Accuracy	Loss
74	0.4	80	0.9

5. ResNet50

The ResNet50 model has 48 convolution layers which include1 MaxPool [9]. Along with this it also includes 1 Average Pool layer in the model [10]. In these, the 2-layers of the block are replaced in the form of 34-layers which also includes the bottom 3-layers of the block. With a stride of size 2, there are a number of different kernels. A convolution model with a kernel size of 7×7 and providing 1 layer to the model. In the next convolution, there are 1×1 and 64 kernels which are followed by a kernel size of 3×3 with 64 kernels. One more convolution has a kernel size of 1x1 with 256 kernels. In order to give, these three layers are repeated three times and achieve an accurate result for the model. These altogether provide 9 layers to the model which is greater when compared to previous convolution.

Table 6. Training and Validation of ResNet 50.

Training		Validation	
Accuracy	Loss	Accuracy	Loss
76	1.6	82	0.6

Next is the kernel size of 1x1 with 128 kernels. After this another kernel is also applied which is of 3×3 size with

128 kernels. At last, a kernel size of 1×1 with 512 is repeated 4 times which provides 12 layers. After this process there is a kernel size of 1×1 with 256 kernels and there are two more kernels. Including the kernel size of 3×3 with 256 and 1×1 with 1024, then this process is repeated 6 times. 18 layers are formed with this technique. Once again, process is performed which includes 1×1 kernel size with 512 kernels with two and more layers. Similarly, a kernel size of 3×3 with 512 kernels and 1×1 with 2048 kernels is performed and this process is also repeated 3 times in order to get the 9 layers. Finally, here an average pool is used at the end which is fully a connected layer and it also contains 1000 nodes with a softmax function which helps in providing 1 layer [11].

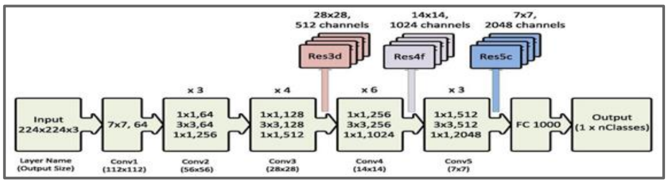

Fig. 8. ResNet 50.

The accuracies obtained by previous authors using Table 7 gives an overview of various pre-trained model types.

Table 7. ResNet 50Training and Validation

Author Details	Method Used	Accuracy Obtained	Remarks
Zhao *et al.* [12]	VGG16	95.24%	Used VGG16 along with backward generative adversarial network (BGAN).
Tekade and Rajeswari [13]	VGG16	95.6%	Used VGG16 along with Lung Nodule Analysis 2016 (LUNA 16)
Pang *et al.* [14]	VGG16	85%	Used for the early-stage detection of lung cancer
Wang *et al.* [15]	VGG16	95.37%	Used VGG16 along with gradient boosting decision tree (GBDT) using chest X-ray 14 dataset.
Present work	VGG16	96%	VGG16 with fine-tuned and augmentation

6. Conclusions

From the above experiment results, it is seen that the performance of ResNet50 proved to give poor accuracy as compared with VGG-16 and Inception-V3. The losses in the ResNet50 pre-trained model seem to have a greater value as compared to the other two models when it is considered at the time of training and validation. Similarly, when training and validating, the inception v3 is creating the fewest losses. but this model unable to deliver satisfactory accuracy results for the testing dataset. Whereas in the case of the VGG-16 model overall performance was found to deliver satisfactory results with good accuracy and minimum losses but the problem of accuracy seems to be very less in this method as well. However, VGG-16 with fine-tuned and image augmentation efficiency enhanced to 96% in the current work in comparision to previous VGG16 based works Usage of Inception V3 and ResNet50 could not obtain the desired accuracy levels for the detection of lung cancers.

Table 8. Training and Validation of ResNet 50.

Sl. No.	Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	VGG Model with Feature Extractor	92	0.2	88	0.6
	VGG Model with Feature Extractor and Image Augmentation	94	0.5	90	0.3
	VGG Model with Fine Tuning and Image Augmentation	96	0.2	93	0.6
2	Inception V3 Model	74	0.4	87	0.9
3	ResNet 50 Model	76	1.2	82	0.5

Due to minimum losses during sing the v2 with Resnet50 is found to have lower accuracy during the training and validation phases. However, this way it is next to impossible to obtain reasonable accuracy while testing the dataset. High conceptual power and higher time is needed for learning entire images while dealing with Inception V3. Whereas in case of ResNet50 with the feedback loop, this model learns for a heavy time and the overall training time for the network may need some weeks of time to train the data. However, on finding the bad training data it delivers bad accuracy for the lung images. Due to this problem ResNet50 generates less accuracy while compiling the lung images.

Acknowledgment

The authors would like to express their gratitude to Annamalai University Management, TN, and Matrusri Engineering College, TS for providing this opportunity to conduct this research with ample support and guidance from the senior professors and experts from industry experts.

References

[1] H. Q. Nguyen, and M. N. Do. "Downsampling of signals on graphs via maximum spanning trees." *IEEE Transactions on Signal Processing* 63, no. 1 (2014): 182-191.

[2] N. Ahmadi, and G. Akbarizadeh. "Hybrid robust iris recognition approach using iris image pre-processing, two-dimensional gabor features and multi-layer perceptron neural network/PSO." *Iet Biometrics* 7, no. 2 (2017): 153-162.

[3] D. Marmanis, M. Datcu, T. Esch, and U. Stilla. "Deep learning earth observation classification using ImageNet pretrained networks." *IEEE Geoscience and Remote Sensing Letters* 13, no. 1 (2015): 105-109.

[4] S. Khalid, T. Khalil, and S. Nasreen. "A survey of feature selection and feature extraction techniques in machine learning." In *2014 science and information conference*, pp. 372-378. IEEE, 2014.

[5] Z. Tüske, M. A. Tahir, R. Schlüter, and H. Ney. "Integrating Gaussian mixtures into deep neural networks: Softmax layer with hidden variables." In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4285-4289. IEEE, 2015.

[6] K. Simonyan, and A. Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.15 56* (2014).

[7] X. Jin, C. Xu, J. Feng, Y. Wei, J. Xiong, and S. Yan. "Deep learning with s-shaped rectified linear activation units." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1. 2016.

[8] A. Odena, C. Olah, and J. Shlens. "Conditional image synthesis with auxiliary classifier gans." In *International conference on machine learning*, pp. 2642-2651. PMLR, 2017.

[9] B. Lakshmi, and S. Parthasarathy. "Human action recognition using median background and max pool convolution with nearest neighbor." *International Journal of Ambient Computing and Intelligence (IJACI)* 10, no. 2 (2019): 34-47.

[10] S. Hershey, S. Chaudhuri, D. P. Ellis, J. F. Gemmeke, A. Jansen, R. C. Moore, M. Plakal, D. Platt, Saurous, R. A., B. Seybold., and Slaney, M. "CNN architectures for large-scale audio classification." In *2017 ieee international conference on acoustics, speech and signal processing (icassp)*, pp. 131-135. IEEE, 2017.

[11] B. Gao, and L. Pavel. "On the properties of the softmax function with application in game theory and reinforcement learning." *arXiv preprint arXiv:1704.00805* (2017).

[12] S. Pang, F. Meng, X. Wang, J. Wang, T. Song, X. Wang, and X. Cheng. "VGG16-T: A Novel Deep Convolutional Neural Network with Boosting to Identify Pathological Type of Lung Cancer in Early Stage by CT Images." *International Journal of Computational Intelligence Systems* 13, no. 1 (2020): 771-780.

[13] D. Zhao, D. Zhu, J. Lu, Y. Luo, and G. Zhang. "Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16." *Symmetry* 10, no. 10 (2018): 519.

[14] R. Tekade, and K. Rajeswari. "Lung cancer detection and classification using deep learning." In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1-5. IEEE, 2018.

[15] M. Wang, D. Li, and L. Tang. "A Lung Image Classification Method: A Classifier Constructed by Combining Improved VGG16 and Gradient Boosting Decision Tree." *Journal of Mechanics in Medicine and Biology* (2021): 2150042.

Madhavi Aluka
Research Scholar
Annamalai University

Dr. Sumathi Ganesan
Ph.D, CSE
Assistant Professor
Annamalai University

Dr. P Vijaya Pal Reddy
Professor,CSE
HOD
Matrusri Engineering College