INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 26-09-2024

Accepted: 05-02-2025

Published: 25-02-2025

Citation: Sridevi C, Kannan M (2025) Convolutional Neural Network Architecture-Inception (GoogLeNet) For Deep Architected Learning-Assisted Lung Cancer Classification in Computed Tomography Images . Indian Journal of Science and Technology 18(7): 504-516. https://doi.org/10.17485/IJST/v18i7.3086

* **Corresponding author**.

sridevijegan@annauniv.edu

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Convolutional Neural Network Architecture-Inception (GoogLeNet) For Deep Architected Learning-Assisted Lung Cancer Classification in Computed Tomography Images

C Sridevi[1]*, M Kannan[2]

1 Associate Professor, Department of Electronics Engineering, MIT Campus of Anna University, Chennai, Tamil Nadu, India
2 Professor, Department of Electronics Engineering, MIT Campus of Anna University, Chennai, Tamil Nadu, India

Abstract

Objectives: The primary objective of this study is to investigate the performance of AlexNet and GoogLeNet architectures in classifying lung cancer images from the LIDC-IDRI dataset. Additionally, the AlexNet architecture is modified to generate three classes (benign, malignant, and non-nodules) using a new dataset. **Methods**: This study utilizes the LIDC-IDRI dataset, consisting of 1018 thoracic CT scans from 1010 patient cases, with annotations from four radiologists. The AlexNet and GoogLeNet architectures are employed for image classification, with the following Hyperparameters: AlexNet uses a learning rate of 0.01 and a dropout of 0.5, while GoogLeNet uses a learning rate of 0.03 and a dropout of 0.4. Both architectures are trained and tested using the LIDC-IDRI dataset. The training accuracy achieved for VGG 16 and GoogLeNet is 0.952 and 0.963 respectively with low log loss. **Findings**: The findings of this study indicate that GoogLeNet outperforms VGG 16 in terms of accuracy, with a classification accuracy of 0.963 compared to 0.952 for VGG 16. Additionally, GoogLeNet demonstrates a higher F1-score (0.946) compared to AlexNet (0.714) and VGG 16 (0.825). The receiver operating characteristic (ROC) curve analysis reveals that GoogLeNet has a higher area under the curve (AUC) value (0.973) compared to AlexNet (0.923) and VGG 16 (0.968). This study demonstrates the effectiveness of deep learning architectures in classifying lung cancer images, with GoogLeNet outperforming AlexNet in terms of accuracy and F1-score. The proposed methodology can be extended to develop a hybrid convolutional neural network for identifying cancer cells in image datasets with improved performance and reduced computational time. **Novelty/application**: Future studies can focus on developing a hybrid convolutional neural network that combines the strengths of AlexNet and GoogLeNet for improved performance and reduced computational time. Additionally, investigating the application of transfer learning and fine-tuning techniques can improve the performance of deep learning architectures in lung cancer image classification.

Keywords: Lung Cancer; CNN Architecture; GoogLeNet; AlexNet; Performance Analysis

1 Introduction

Radiologists utilize computer tomography (CT) scans to detect and monitor cancer within the body[1]. However, the subjective nature of human interpretation, coupled with visual complexity and variability among interpreters, restricts the accuracy of image analysis[1]. To address this challenge, image segmentation and radiomic feature extraction from CT scans have emerged as crucial components of image processing[1]. Furthermore, deep learning techniques have revolutionized medical imaging by providing innovative and accurate solutions, thereby establishing themselves as a vital tool for future healthcare applications[2]. A comprehensive survey conducted by Kumar[2] highlights the efficacy of shallow and deep learning techniques in medical imaging. The study reveals that deep learning techniques exhibit a higher accuracy rate in detection compared to other methods[2]. Moreover, a fully automated system for chest disease detection, developed by Imran[3], utilizes a modified Nested UNet architecture to achieve an accuracy of 92.86%. This system integrates classification and segmentation tasks into a single model, thereby ensuring a lightweight design that minimizes overfitting[3]. Deep learning techniques have also been employed for lung cancer detection and stage classification, yielding superior results[4]. A study by Deepa[4] explores the application of different deep learning techniques, including the utilization of pre-trained models such as AlexNet, ResNet18, GoogleNet, and ResNet50. The performance of these models is evaluated using metrics such as confusion matrix, precision, recall, specificity, and f1-score[4]. Accurate segmentation of lung tumors is essential for effective cancer detection[5]. To achieve this, a 3D-convolutional neural network (3D-CNN) is applied to separate cancerous regions, which are then integrated with a Markov model-based active contour deformable model to enhance the average Dice similarity coefficient[5]. A neural network model is also proposed for identifying cancer cells in CT images and recognizing issues in therapeutic imaging applications[6].

A framework originating from this research encompasses the application of deep neural networks and AI for lung cancer identification[6]. The framework comprises multiple steps, including image collection, pre-preparation, pixel enhancement, segmentation, and feature extraction[6]. A study by[7] demonstrates the efficacy of a trained deep learning model in achieving accurate lung segmentation, with an accuracy of 99% and an IOU index of 0.99. To address the challenges of lung cancer detection, this study proposes an automatic lung nodule classification method utilizing deep learning techniques[8]. The results show that transfer learning with MobileNet achieves the highest accuracy rate of 99.11%[8]. Consequently, the proposed model can classify patient CT images into three categories: Normal, Pneumonia, and COVID, with a guaranteed high level of accuracy[9]. The Data Classification technique and View (DCV) are crucial components of the system, ensuring better lung cancer classification results[10]. The proposed work utilizes image data with varying dimensionalities as input to the deep learning-based classifier, providing lung cancer classification for radiologists to facilitate early diagnosis[10]. A deep modified CNN network is proposed, leveraging three pre-trained models – Densenet169, MobileNetV2, and Resnet50V2 – to enhance lung cancer classification performance based on the IQ-OTH/NCCD Lung Cancer dataset[11].

The Conv-Unet design adds convolutional layers to the standard U-Net architecture, capturing intricate patterns and boundaries of lung nodules for more accurate segmentation[12]. Categorization of segmented lung nodules is performed using a CNN network on the LIDC-IDRI and LUNA16 datasets[12]. This research work develops a novel algorithm, MediNet, for segmenting cancer-affected regions in CT images, achieving 97% accuracy [13]. Lung segmentation poses a challenge due to variability caused by age, gender, or health status, becoming even more complex with the presence of external objects [14]. As a result, accurate lung field segmentation is regarded as a critical task in medical image analysis [14]. According to the World Health Organization, lung cancer accounted for approximately 1.8 million deaths in 2020, emphasizing the need for early diagnosis and effective screening methods [15]. This paper aims to detect and classify lung cancer using deep learning, providing encouraging results with high accuracy rates for medical assistance [15]. Automatic detection and classification of COVID-19 from chest X-ray images using computer vision technology offer a valuable complement to traditional detection methods [16].

2 Methodology

2.1 Proposed Method

The flow diagram (Figure 1) describes the methodology used for the classification of lung cancer using deep neural networks. The Dataset contains three classes of Lung cancer sorts which are Adenocarcinoma, Huge cell carcinoma, Squamous cell carcinoma, and one organizer for the typical cell. All pictures are in PNG and arranged to fit the demonstration. The dataset contains up to 1100 pictures speaking to the CT check cuts of all the classes.

The inclusion of a specific element named Inception is the principal advancement in the GoogLeNet architecture. The Inception module's fundamental concept is the parallel execution of numerous operations (pooling, convolution) with various filter sizes (3x3, 5x5, etc.). The starting input is a tensor of 64 feature maps, each of which has dimensions of 32x32 and represents the stack of feature maps produced by a previous layer. The GoogleNet Architecture has a total of 22 layers, including 27 pooling layers. There are a total of 9 linearly stacked inception components. The global average pooling layer is connected to the ends of the inception modules. Distributed machine learning systems with significant model and data parallelism are used to train GoogleNet. The learning rate was reduced by 4% per eight epochs during training using asynchronous stochastic gradient descent with a velocity of 0.9. Object detection, image classification, and other computer vision applications use GoogleNet. On image classification, roughly 18% of GoogleNet apps are based, and object detection and quantization, about 10–15% of GoogleNet applications.

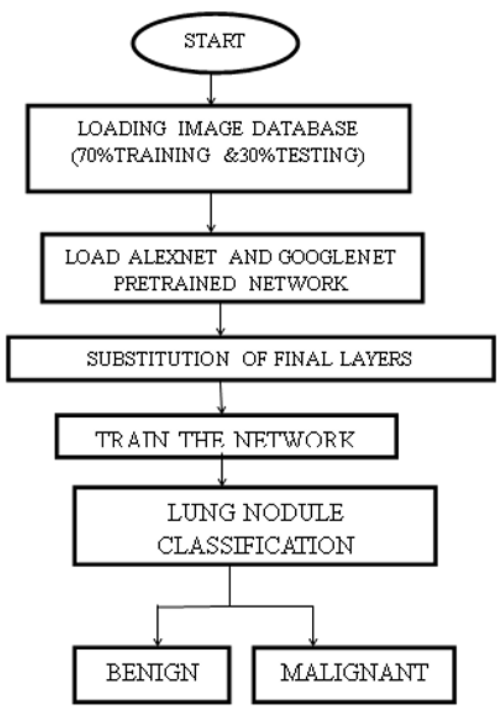

Fig 1. Overall Process Flow Diagram

Preprocessing

In the Preprocessing step of the input lung CT images, as the AlexNet architecture demands RGB images as the input, the initial step is to convert the given input images into RGB of dimension 227 X 227 X 3. The preprocessed data is trained with the ConvNet of 20 epochs using 30% of the images for training purposes and 70% of the images for testing purposes with the initial learning rate set at 0.001.

In this proposed architecture, the AlexNet consists of eight layers, 5 convolutional layers and three fully connected layers. In this architecture, the models introduced are the maxpooling and the ReLU activation function. The input images for AlexNet are of the size 227 X 227 X 3 where the maxpooling layer is designed with a 3 X 3 filter with stride 2 and a convolution layer of 11 X 11 with stride of 4. The pre-trained AlexNet model depicts the modification in the network where the elimination of the fully connected layers was done, and the network generates the classified two classes. In this step, the class score of 1000 has been changed to 2 i.e. benign or malignant. The proposed architecture comprises 5 convolutional layers, where the first 2 convolution layers are followed by the maxpooling layer and the second maxpooling layer is designated after the third convolutional layer. The hyper parameters generated in the AlexNet Architecture have a fully connected layer where the output size is set to 2 which indicates two classes and the initial weight learn rate factor is set to 10. Figure 4**(a) & (b)** illustrate the Classification results of the cancer nodules as benign and malignant using AlexNet Architecture.

Three different filters of size 1x1, 3x3, and 5x5 are used in the GoogLeNet architecture, which enhances the robustness of the network by combining the features obtained as output. The architecture comprises 22 layers and also minimizes the number of parameters used in the network from 60 million to 4 million. The major feature of the GoogLeNet architecture is the introduction of the 1x1 convolution for dimensionality reduction. The major advantageous feature of this network is that it chooses the best weight from the network during the training process itself and selects the prominent features for the classification purpose[17]. This leads to the formation of a deeper network that consists of multiple inception modules, enabling to acheive good accuracy in classification. The pre-trained model of GoogLeNet consists of the actual layers of the loss classifier, which is replaced with the fully connected layers and the classification layers is replaced by the two classes i.e. benign and malignant. The hyperparameters of a fully connected layer of GoogLeNet, where the output size is set to 2 which indicates two classes and initial weight learn rate factor is set to 10.**Figure 4(a) & (c)** illustrate the Classification results of the cancer nodules as benign and malignant using GoogLeNet Architecture.

Transfer Learning Using Alexnet

The network training for the applications involving large datasets is a monotonous process and involves more amount of time. In order to avoid the unwanted waste of time and use of network modules, pre-trained networks are engaged in the applications for extracting features and involve large datasets. This may be termed as transfer learning[12] and the network involved in this process is the AlexNet. The highlighted features of the transfer learning process are enlisted below: (a) ConvNet as a fixed feature extractor, (b) Fine- tuning the ConvNet, (c) Pre- trained models.

Negative transfer refers to the reduction of accuracy of a deep learning model after retraining (biologically, this refers to the interference of previous knowledge with new learning). The Challenges associated with transfer learning in Natural and Biomedical Images are:

- Availability of Database is the biggest problem in Medical Image Processing.
- Complexity involved in the Image Texture.
- In Medical images, recognition is not sufficient, prediction of future values is a mandate.

One of the biggest difficulties in processing medical images is the non-standard image acquisition. To guarantee that the deep learning network produces a robust algorithm, the dataset must be larger the more variability there is in the data. The vast collection of images, the high complexity of the data, and the dearth of labeled data pose the biggest difficulties in image categorization. With many pixels, images may grow to be rather huge. Each image may contain high-dimensional data with a wide variety of attributes. To perform better than other strategies, it needs a tremendous amount of information. Because to sophisticated data models, training is quite costly. Deep learning also needs hundreds of workstations and pricey GPUs. The cost to users goes up as a result.

Proposed Architecture

Deep learning is a process involving large amount of data for training process and this condition of data handling may be termed as Data Hungry. In order to overcome the shortage of data, the available data are annotated and taken from various angles, illumination conditions etc., this process may be termed data augmentation, where the data can be multiplied and taken for the training process. The data are multiplied in terms of a mirror image of the actual image or intensity varied image or distorted image of the original image. The problems faced in the deep learning models are the overfitting problem and the lack of generalization of the model. In order to avoid the overfitting problems in the deep learning approaches, gradient descent techniques can be employed to overcome the errors like fading of features that occurred during the training process.

The architectures used for comparison in the proposed work are AlexNet and GoogLeNet. The AlexNet architecture comprises five convolutional layers, the maxpooling layers and the softmax layer [13]. In the architecture, the two convolutional layers are placed before the maxpooling layer and the three convolutional layers are placed after the maxpooling layer. In the architecture, the last layer is the softmax which is placed before the fully connected layers. The softmax layer differentiates the preprocessed image samples into two classes either it is malignant or benign. The convolution operation of the preprocessed input images was carried out by the convolutional layer which forms the basic module of the ConvNet. The maximum value from each of the previous layer clusters is used by the maxpooling layer, which is the consecutive layer of the convolution layer. The maxpooling layer minimizes the spatial dimension of the activation maps. The final classification is done in the fully connected layer which evaluates the class score of the algorithm also the layer converts the pixel value into the class score. The dropouts and the loss function reduce the over fitting problem in the fully connected layer. In this proposed architecture, softmax is the loss function applied.

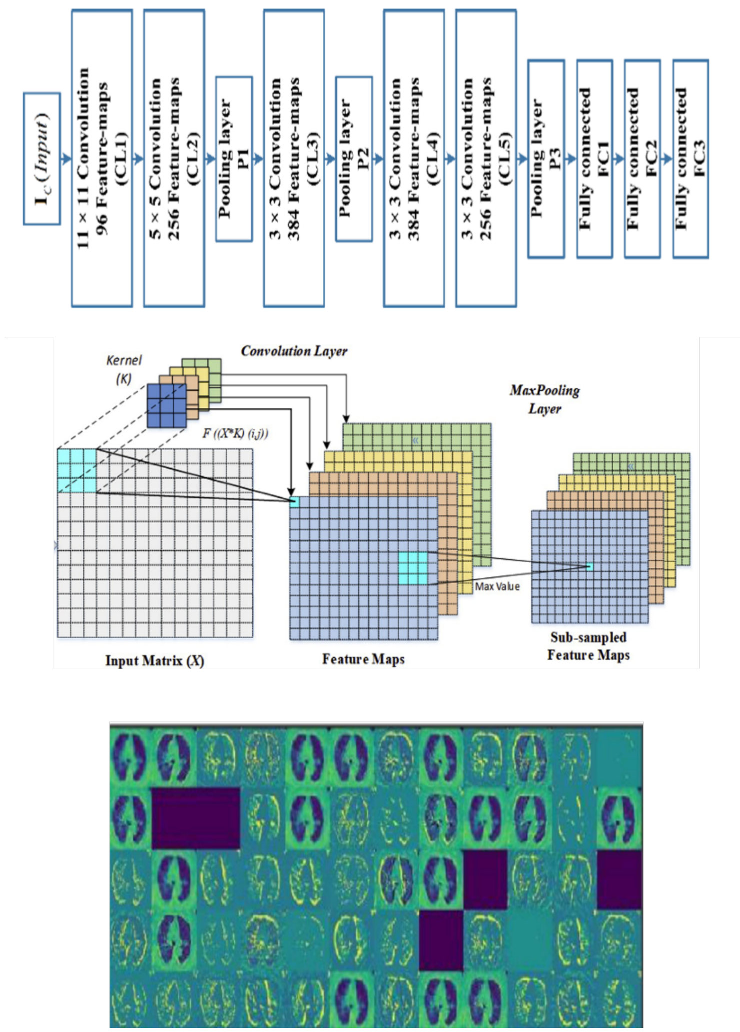

Fig 2. Alexnet Architecture (Four Convolution Layers And Three Fully Connected Layers), Feature Extraction Phase, Convolution Layer Output

Figure 2 explains the extraction of features from the images which could reveal shape, edges, corners, intensity variations etc. it is useful when we have a large dataset and need to reduce the number of resources without losing any important or relevant information. It also illustrates the features extracted after the convolution layer and also depicts the features that are extracted after the max pooling layer. These features hold subtle changes in an image without loss of any pixel information. The feature extraction process helps to reduce the amount off redundant data from the datasets.

Procedures Involved In The Alexnet Architecture

The MATLAB function "splitEachLabel' divides 70% of the images for training and the remaining 30% for testing. The new set of Image datasets is loaded by unzipping the 1018 images. After loading the images, the Pre-trained AlexNet network is loaded, which reveals the functionality of the network layers by exhibiting the weight, bias, stride, and padding used in the convolutional ReLU and maxpooling layers[18]. The architecture and layer information of the AlexNet and the algorithmic flow of the network by showing the quantitative parameters. By loading the AlexNet pre-trained network. It will display the information of the network layer. The third step involves the substitution of the three layers with fully connected layers that include the classification output layer and softmax layer. The training process involved in the AlexNet Architecture involves the MATLAB function 'trainNetwork (imds,layers,options)' is used for image classification. The imds - accumulate the input images, layers - specify network architecture, options - "InitialLearnRate, MiniBatchsize, InitialLearnRate, MaxEpochs, ValidationData, ValidationFrequency, plots, training progress" are the commands used in the algorithm to detect the cancerous cells present in the lungs. The outputs are obtained from the loaded input images illustrating the two classes Benign 88.5% and Malignant 90.2%, shown in Figure 6 for the AlexNet architecture.

PSEUDO CODE FOR ALEXNET:

Input: Lung images (LIDC-IDRI) Dataset
Output: classified Lung cancer images
1: INITIALIZATION:
Optimizer = SGDM
MINIBATCHSIZE = 10
MAXEPOCHS = 6
INITIALLEARN RATE = 1e-4
ValidationFrequency=3
Verbose= false
2. Split labels 70% training and 30% testing sets
3. Open deep network designer app
4. From the layer library connect input layer, convolutional, pool layers
5. Replace final layers such as fully connected layer and classification layer to learn feature specific to database
6. Set the output class size = 2
7. Set the weight bias factor =10
8. Load pretrained network from workspace
9. Analyze the Alexnet network architecture
10. Create and Compile the model and export the network
11. Train the Network
12. Evaluate and access the training results
13. Test the model and visualize the results

GOOGLENET

GoogLeNet Architecture – Inception – IV focuses on high accuracy with reduced with reduced computational cost. In the convolutional neural network, the inception block incorporates Split, Transform and Merge for the multi- scale convolutional transformations[14]. Figure 3 displays the inception block architecture involving split, transform and merge concepts. As projected in the Network in Network (NIN) architecture, the convolutional layers used in the conventional methods are replaced by smaller blocks in the GoogLeNet architecture. These smaller blocks comprise filters of different sizes i.e. 1x1, 3x3, 5x5, which include both the fine and coarse grain levels in order to acquire spatial information at different scales. The GoogLeNet architecture mainly focuses on the toughness faced in learning the same category of images with different resolutions and also standardizes the computational complexities of the large kernel sizes, by appending a 1x1 convolutional filter. In addition to that feature, the GoogLeNet also includes a sparse connection, in order to overcome the problem of redundant information and reduced cost of the architecture. Also, the architecture also has the advantageous feature of introducing the global averaging pooling at the final layer instead of using 25 fully connected layers for reducing connection density.

These advancements in the network architecture show a massive variation in the significant decrease of the parameters from 138 million to 4 million. The other regulatory factors highlighting the network are batch normalization and the usage of RmsProp as an optimizer. The convergence rate of the network is accelerated by the introduction of auxiliary learners in the network. The major bottleneck of the GoogLeNet architecture is the heterogeneous topology which has to be customized from nodule to nodule and also the architecture randomly reduces the feature space which may lead to the loss of valid information. Figure 3**(a)** illustrates the features extracted after the convolution layer and also depicts the features that are extracted after the max pooling layer. These features hold subtle changes in an image without loss of any pixel information. The feature extraction process helps to reduce the amount off redundant data from the datasets.

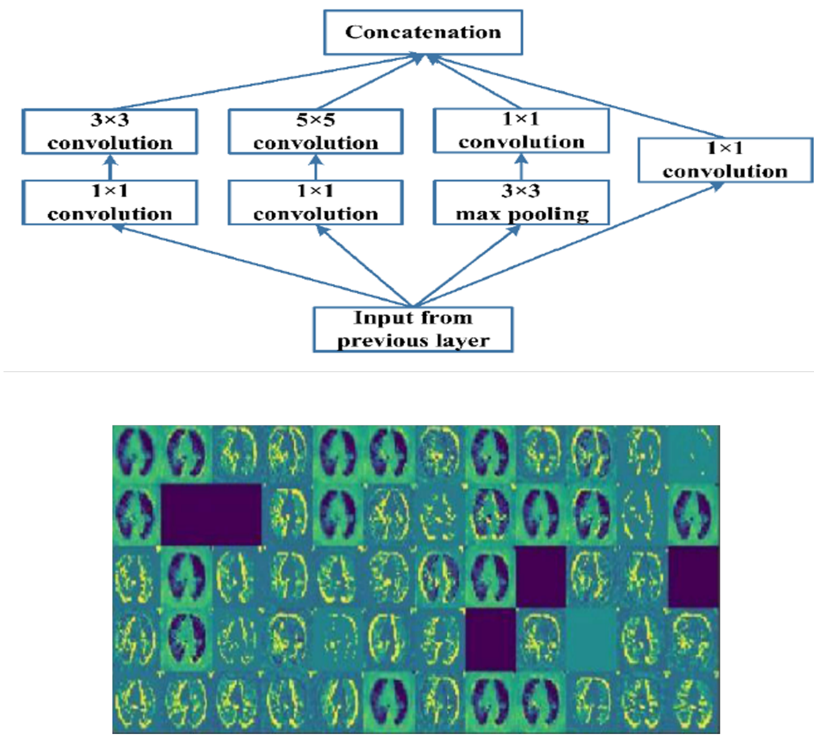

Fig 3. Split, Transform And Merge Concept (Inception), Maxpooling Layer Output

PROCEDURES INVOLVED IN THE GOOGLENET ARCHITECTURE

In the GoogLeNet Architecture, the classification of lung cancer involves the loading of the images from the dataset for the training and validation process, in which 70% of the images were involved in the training process and 30% of the images were involved in the testing process. Figure 3 shows the deep network designer output of the GoogLeNet. Once the network is loaded, "analyzeNetwork" command aids in displaying information on the architectural parameters involved in the proposed network.

The data's related to the weights, and bias of the convolution network of the GoogLeNet architecture and also shows the extracted features in Figure 3**(a)**. The analysis of the classification layers is done by extracting the features from the input images. The extracted features are then combined by the "loss3- classifier" and "output" layers into loss value, predicated label, and class probabilities. In order to retrain the new dataset, set the initial layers to zero. The network training speed can be enhanced by significantly freezing the layers. The size of the images to be trained should be 223x223x3 size.

The images in the selected datasets may vary in terms of size. The training process is restored using the augmented image data store to preserve the size of the image. Augmentation of data focuses on training the network and also avoids overfitting and, at the same time retains the characteristics of the training images. The classification of the authorized images using predicted labels and also forecast the probabilities of the images. Fig. 9, shows validation images along with predicted labels and prediction probability. Using the GoogLeNet architecture, the accuracy for the benign is 99.8%, and for Malignant 95.7%.

PSEUDO CODE FOR GOOGLENET
Input: Lung images (LIDC-IDRI) Dataset
Output: classified Lung cancer images
1. INITIALIZATION:
Optimizer = sgdm
MiniBatchSize = 10
MaxEpochs=6
InitialLearnRate=0.001
ValidationFrequency=3
Verbose= false
2. Split labels 70% training and 30% testing sets
3. Import new dataset
4. Open deep network designer app
5. From the layer library connect input layer, convolutional, pool layers
6. Replace final layers such as loss classifier layer to identify the features specific to the dataset using fully connected layers.
7. Set the output class size=2
8. Set the weight bias factor =10
9. Freeze the network weights
10. Load pre trained network from workspace
11. Analyze the GOOGLENET network architecture
12. Create and Compile the model and export the network
13. Train the Network
14. Evaluate and access the training results
15. Classify the lung cancer with predicted probabilities

VGG16

VGG16, a convolutional neural network (CNN) architecture, achieved victory in the 2014 ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). It remains a highly influential vision model. Its defining characteristic is the consistent use of small 3x3 convolutional filters with a stride of 1 and "same" padding, coupled with 2x2 max pooling layers with a stride of 2. This uniform configuration of convolutional and max pooling layers is maintained throughout the entire architecture. The network concludes with two fully connected layers followed by a softmax output layer.

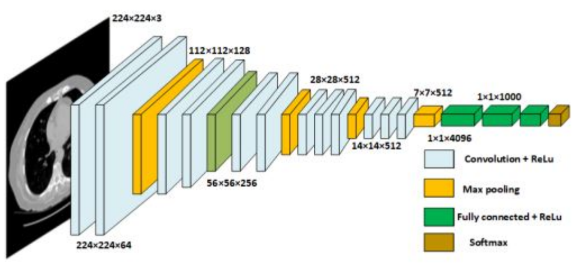

Fig 4. G eneral network architecture of VGG 16

Figure 4 shows the architecture of the pre-trained VGG16 model. VGG16, initially trained on a large image dataset, utilizes the weights learned during that process. The initial layers of the model extract basic image features, the middle layers detect shapes, and the top layers perform complex feature extraction and classification. Because lung cancer detection involves classifying images into two categories (cancerous and non-cancerous), the original output layer of VGG16 is replaced with a two-node dense layer. By freezing the lower layers and fine-tuning only the upper layers, the amount of required training data is reduced, and training time is significantly faster compared to training a CNN from scratch.

3 Results and Discussion

The performance of the model can be evaluated using the validation set of images, whereas the training set of images is used to train the model. The validation of the architecture is done with the Monte Carlo cross validation, multiple random sets of data are provided for processing both in the training and in the validation part. In this validation algorithm, each random split is adequate for the training data and the algorithm accuracy can be evaluated using the validation dataset. The same procedure is repeated for the other random sets and the results are averaged.

For calculating result performance metrics such as binary accuracy and log loss of AlexNet and GoogLeNet architecture are calculated using the equation (1) and (2) respectively. An accuracy metric in this algorithm defines the performance of the network in an understandable manner. The model defines the accuracy of the architecture as a measure of how efficiently the algorithm predicts the results compared to the desired data. The optimization of the machine learning algorithm is defined using the metric log loss function. The elucidation of the log loss calculated on the training and the validation images is based on how the network behaves for the training and testing set of images. Log loss value interprets the model behaviour after each iteration of optimization.

$$Accuracy = \frac{Number\ of\ Truly\ Classified\ samples}{Number\ of\ samples} \tag{1}$$

$$Hp(q) = \frac{-1}{N} yi \left\{ \log\left[p(yi)\right] + (1 - yi) \,.\, \log[(1 - p(yi))] \right\} \tag{2}$$

”Dropout” is used to define the number of layers dropped out in the layer of the network in both the hidden and the visible layers. It can also be defined as deleting the unwanted neurons during the training phase which is selected at random. During the training phase, the nodes dropped out from the network can be given with the probability (1-p) and the neurons present in the network are given by the probability p, which results in the minimized network for the process, which also eliminates the inputs and the outputs of the dropout nodes.

(a) (b) (c)

Fig 5. Classification Output (a) Benign (b) Malignant-AlexNet (c) Malignant- GoogLeNet

GoogleNet has 144 layers which increases the accuracy but it can increase the overfitting problem. The overfitting problem can be overwhelmed by data augmentation. The data augmentation includes Scale Range which flips the images vertically and randomly transforms them to 30 pixels and it scale them likely to 10% both horizontally and vertically.

Table 1. Performance Comparison of Existing and Proposed Architecture

Models	Accuracy	F1 Score	Recall	Precision	Area under the Curve (AUC)
Traditional CNN	0.856	0.648	0.783	0.726	0.823
AlexNet	0.903	0.714	0.813	0.795	0.923
GoogLeNet	**0.963**	**0.946**	**1.000**	**0.898**	**0.973**
VGG 16	0.952	0.825	0.965	0.887	0.968

The Table 1 shows that GoogLeNet performs better than AlexNet and Traditional CNN Algorithm. The higher the precision and recall, the higher the F1- Score. The Closer the value to 1, the higher the model performance. By dividing the genuine positives by any positive predictions, precision is determined. By dividing the real positives by everything else that should have been projected as positive, recall (also known as the real Positive Rate) is obtained.

The use of rectified linear unit (ReLU) activation functions, which helped to mitigate the vanishing gradient issue; (i) a deep architecture with three fully connected layers after five convolutional layers; and (ii) a regularization technique called dropout, which prevented overfitting by randomly removing neurons during training, are some of AlexNet's key classification features. AlexNet can be tuned to generate pixel-wise predictions for segmentation tasks, despite the fact that its extremely shallow architecture may limit its performance.

Building on AlexNet's success, GoogleNet proposes a number of advancements that significantly enhance categorization performance. Among GoogleNet's salient characteristics are: (i) a deeper architecture comprising 22 layers, which combines convolutional and inception layers; (ii) the utilization of inception modules, which enabled a notable expansion in depth and width while maintaining a constant computational cost; and (iii) a global average pooling layer, which decreased the number of parameters and enhanced the model's translation invariance. Fully convolutional networks (FCNs), which provide effective pixel-wise predictions, are one method for adapting GoogleNet for segmentation tasks.

The sample 50 images are selected and analyzed for training and testing purposes using the concept of confusion matrix are illustrated in Figure 6 & Figure 7.

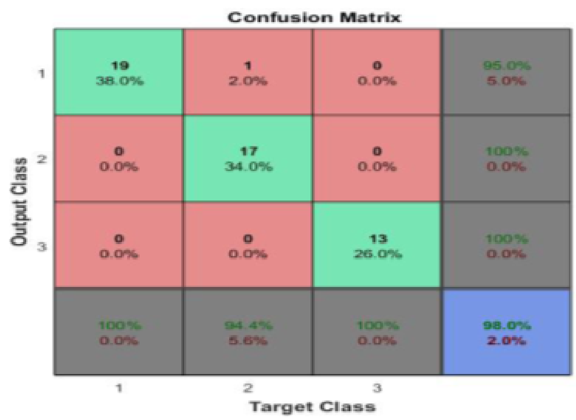

Fig 6. Confusion Matrix For 50 Sample Training Images (Googlenet)

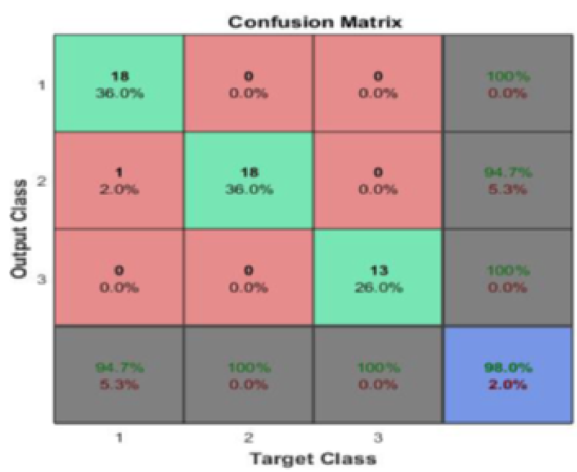

Fig 7. Confusion Matrix For 50 Sample Test Images (GoogLeNet)

A misclassification is shown in the plot between the predicted and the original image. The percentage of images classified as Benign and malignant is also illustrated in the plot. After obtaining the confusion plot of the training set images, the testing set images are taken for the analysis and a misclassification is shown in the second row of the plot. Similarly, the confusion matrix is built for the 105 training samples of both the classes i.e. Benign or Malignant shown in Figure 8.

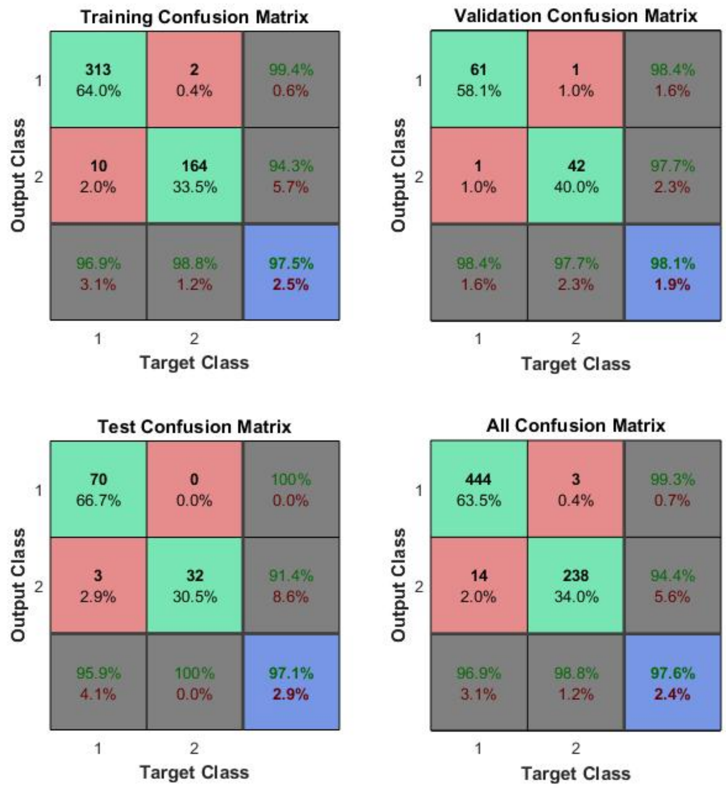

Fig 8. Confusion Matrix For 105 Sample Training Images (Googlenet)

Cross-validation is a statistical technique that involves partitioning the data into subsets, training the data on a subset and using the other subset to evaluate the model's performance. To reduce variability, multiple rounds of cross-validation with different subsets from the same data are engaged. The validation results from these multiple rounds to come up with an estimate of the model's predictive performance are combined to form optimal results.

Cross-validation will give a more accurate estimate of a model's performance. In Leave One Out Cross Validation (LOOCV) the data set is divided into two parts. In one part, single observations are made and these observations are treated as Test data and in the other part, all the observations from the dataset are treated as training data. If there is a data set with n observations then training data contains (n-1) observation and test data contains 1 observation. This process is iterated for each data point. Repeating this process n times generates n times Mean Square Error (MSE). GoogLeNet Architecture has a feature of including the Validation accuracy and loss calculation for both the Training and Test data.

GoogLeNet classifies the benign and the malignant images during the classification process more effectively than the AlexNet architecture in terms of the number of layers and the choice of minimal number of parameters. GoogLeNet architecture possesses an efficient performance in terms of metrics like speed, accuracy, dropout, and the initial learning rate. The future scope of this work may be extended as a hybrid convolutional neural network, which identifies the cancer cells in the image dataset at a lesser time duration and yields better performance. The initial learning rate for VGG 16 and GoogLeNet is 0.01 and 0.03 respectively.

The dropout values for AlexNet and GoogLeNet is 0.5 and 0.4 respectively. The training accuracy achieved for VGG 16 and GoogLeNet is 0.952 and 0.963 respectively with low log loss. The more the number of layers in the architecture, the higher the accuracy of the architecture. The testing stage of both the VGG 16 and the GoogleNet Architecture shows a classification index of 90.2% and 95.7% for malignancy identification respectively.

The Lung Cancer Segmentation methodology adopted would offer a more thorough grasp of the model's advantages and disadvantages by including this more thorough examination of instances that were incorrectly classified. In addition to increasing the work's scientific worth, this would lead to future research efforts to increase the precision and resilience of lung cancer segmentation algorithms.

This technique can be applied to large datasets in the future to increase accuracy using machine learning. Furthermore, compatibility with images from other modalities that were obtained using other image acquisition techniques can be guaranteed by utilizing this method. A hybrid convolutional neural network, which performs better and can identify the cancer cells in the image collection faster, might potentially be developed from the work. Future studies may include identifying Multiview techniques to increase the accuracy of the segmentation and detection algorithms. Furthermore, the study can be extended to derive more specific 3D models for the lung and pulmonary nodules for better analysis and assistance. The various ideal parameters are examined for the 3D reconstruction of the lung and the pulmonary nodules.

4 Conclusion

The AlexNet, VGG 16, and GoogLeNet architectures are used in this proposed work to analyze deep learning models and extract features for the classification of lung nodules from the obtained LIDC-IDRI dataset. Classification of lung nodules is done into two groups: cancerous and non-cancerous. Deep learning architectures have recently achieved state-of-the-art outcomes in image classification applications. The classification of benign or malignant images is the problem to be tackled in this suggested effort, and the precision of the data produced is crucial and required.

In order to classify the acquired images as either benign or malignant, it is also important to get the macroscopic features of the images. In terms of the number of layers and the selection of the fewest number of parameters, it is clearly demonstrated that GoogLeNet performs better than AlexNet in classifying both benign and malignant images during the classification process. When it comes to parameters like speed, accuracy, dropout, and initial learning rate, the GoogLeNet architecture performs efficiently.

This work's future scope could be expanded to include a hybrid convolutional neural network, which performs better and recognizes cancer cells in picture datasets in less time. GoogLeNet and AlexNet have started learning rates of 0.03 and 0.01 respectively. GoogLeNet and AlexNet have dropout levels of 0.4 and 0.5 respectively. With little log loss, AlexNet and GoogleNet have training accuracies of 0.90 and 0.95, respectively. The precision of the architecture increases with the number of layers in the design. In terms of identifying malignancy, the VGG 16 and GoogLeNet Architecture testing stages display classification indices of 90.2% and 98.1%, respectively.

In the future, this method can be used on big datasets to achieve more accuracy through machine learning. Additionally, this approach can be used to ensure compatibility with images from other modalities that were acquired using various image acquisition techniques. The work may also be expanded into a hybrid convolutional neural network, which performs better and recognizes the cancer cells in the image collection in less time. Finding Multiview techniques to improve the precision of the detection algorithms and segmentation may also be part of it.

Additionally, for improved analysis and support, the work can be expanded to extract more specialized 3D models for the lung and pulmonary nodules. For the 3D reconstruction of the lung and the pulmonary nodules, the different optimal parameters are investigated. The classification of benign and malignant lung nodules and the decrease in false positives can guide future research. The fact that malignant nodules enlarge more quickly than benign ones is another key element that is ignored. A patient's sequential CT scan data may be used as input during the training phase to analyze the likelihood of malignancy. The classification of the lung nodules as malignant or benign with the help of deep learning network can be further analyzed using the Hybridized Heuristic Mathematical Model for effectively predicting the lung cancer nodules at an earlier stage.

Declaration of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval

This article does not contain any studies with animals performed by any of the authors.

Informed consent

This study includes images from publicly available databases. The Cancer Imaging Archive (TCIA) Datasets have been properly cited.

References

1) Bhattacharjee R, K SD, K S, T D. Deep Neural Network Based Automatic Detection and Classification of Lung Nodules from CT images. In: 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON). IEEE. 2022;p. 1–4. Available from: https://ieeexplore.ieee.org/document/10084070.
2) Swain AK, Swetapadma A, Rout JK, Balabantaray BK. A Review on Lung Cancer Detection and Classification using Shallow Learning and Deep Learning. In: 2023 OITS International Conference on Information Technology (OCIT). 2023;p. 570–573. Available from: https://www.semanticscholar.org/paper/A-Review-on-Lung-Cancer-Detection-and-using-Shallow-Swain-Swetapadma/4116c085b5f12c7062b2f3c9114a2da8be36483b.
3) Imran MH, Shill PC. Automatic Classification and Localization of Lung Diseases from Chest X-Ray Images using Deep Learning. In: 2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET. 2023;p. 1–6. Available from: https://www.researchgate.net/publication/378090253_Automatic_Classification_and_Localization_of_Lung_Diseases_from_Chest_X-Ray_Images_using_Deep_Learning.
4) Deepa V, Fathimal PM. Lung cancer prediction and Stage classification in CT Scans Using Convolution Neural Networks -A Deep learning Model. In: 2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI). IEEE. 2022;p. 1–5. Available from: https://ieeexplore.ieee.org/document/10028880.
5) Shetty MV, Jayadevappa D, Achar S. Automatic Segmentation Methods for Lung Cancer Using Deformable Models and Deep Learning. In: 2023 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT). IEEE. 2023;p. 1–5. Available from: https://ieeexplore.ieee.org/document/10234825.
6) Rao BD, Arshad M. Early Detection of lung Cancer Using Machine Learning Technique. In: 2023 International Conference on Computer Communication and Informatics (ICCCI). 2023;p. 1–5. Available from: https://ieeexplore.ieee.org/document/10128389.
7) N K, Ramyea R, R E, Josika T, Jothish M, Kirithick S. Segmentation and Classification of Acute Respiratory Distress Syndrome (ARDS) using Deep Learning. In: 2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC). IEEE. 2024;p. 1–6. Available from: https://ieeexplore.ieee.org/document/10593893.
8) Ghosh R, Ahamed A, Sadhukhan B, Das N. Lung Nodule Classification Using MobileNet Transfer Learning. In: 2023 9th International Conference on Smart Computing and Communications (ICSCC). IEEE. 2023;p. 290–295. Available from: https://ieeexplore.ieee.org/document/10335043.
9) Wang S, Yang G. A Novel Automated Classification and Segmentation for COVID-19 using 3D CT Scans. In: 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS). IEEE. 2022;p. 1–5. Available from: https://ieeexplore.ieee.org/document/10052819.
10) Tiwari S, Abdullah SH, Mubasher R, Alsadoon A, Prasad PWC. DCV: A Taxonomy on Deep Learning Based Lung Cancer Classification. In: 2021 6th International Conference on Innovative Technology in Intelligent System and Industrial Applications (CITISIA). IEEE. 2021;p. 1–10. Available from: https://ieeexplore.ieee.org/document/9719963.
11) Benamara Z, Zehani S, Zitouni A. The effect of fully connected layers in different pre-trained CNN architectures on the enhancement of lung cancer classification. In: 2024 8th International Conference on Image and Signal Processing and their Applications (ISPA). 2024;p. 1–6. Available from: https://www.semanticscholar.org/paper/The-effect-of-fully-connected-layers-in-different-Benamara-Zehani/89f8bf6983eea5ddd22663379caa4d01c05785f8.
12) V A, R MS, B AK, G JM, Kannadaguli P. Lung Nodule Segmentation and Classification Using Conv-Unet Based Deep Learning. In: 2023 International Conference on Network, Multimedia and Information Technology (NMITCON). IEEE. 2023;p. 1–6. Available from: https://ieeexplore.ieee.org/abstract/document/10276037.
13) Divya M, Sathya S. Novel Lung Cancer Classification and Segmentation Algorithm with Web UI Interface. In: 2022 6th International Conference on Electronics, Communication and Aerospace Technology. IEEE. 2022;p. 1–9. Available from: https://ieeexplore.ieee.org/document/10009080.
14) Penugonda BS, Koganti A, Unnam A, Chinnadurai S. Automated Lung Size Estimation in Chest X-Ray Images Using deep learning. In: 2023 IEEE 20th India Council International Conference (INDICON). IEEE. 2023;p. 863–870. Available from: https://ieeexplore.ieee.org/document/10440878.
15) Maalem S, Bouhamed MM, Gasmi M. A deep-based compound model for lung cancer detection. In: 2022 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS). IEEE. 2022;p. 1–4. Available from: https://ieeexplore.ieee.org/document/9946875.
16) Hasan MJ, Alom MS, Ali MS. Deep Learning based Detection and Segmentation of COVID-19 & Pneumonia on Chest X-ray Image. In: 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). IEEE. 2021;p. 210–214. Available from: https://ieeexplore.ieee.org/abstract/document/9396878.
17) V B, J J, R M, Subramaniyam K. Lung Cancer Detection by Artificial Intelligence Using Deep Learning Techniques. In: 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN). 2023;p. 831–839. Available from: https://ieeexplore.ieee.org/abstract/document/10266283.
18) Jain H, Kumar I, Jain K, Porwal IN, Singh T, Mohd N. Revealing Multiple Lung Condition Using Deep Neural Network Models from X-RAY Images. In: 2023 IEEE World Conference on Applied Intelligence and Computing (AIC). 2023;p. 728–734. Available from: https://ieeexplore.ieee.org/document/10263982.