SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

GAN-Based Image Segmentation and Classification Using Vgg16 for Prediction of Lung Cancer

Vishnu Priyan Swaminathan[1,*], Ramesh Balasubramani[2], S. Parvathavarthini[3], Vidhya Gopal[4], Kanagaselvam Raju[5], Tamil Selvan Sivalingam[6], Sounder Rajan Thennarasu[7]

[1] Department of Biomedical Engineering, Kings Engineering College, Chennai – 602117, India
[2] Department of ECE, Annapoorana Engineering College, Salem 636308, India
[3] Kongu Engineering College, Perundurai, Tamilnadu, India – 638052, India
[4] Computer Science and Engineering, Sona College of Technology, Junction Main Road, Salem, India
[5] Department of Science and Humanities, Nandha College of Technology, Erode, 638052, Tamil Nadu, India
[6] Department of Computer Science and Design, Erode Sengunthar Engineering College, Thudupathi, Perundurai, Erode, India
[7] Department of AI and DS, Erode Sengunthar Engineering College, Thudupathi, Perundurai, Erode – 638057, India

ARTICLE INFO

Article history:
Received 11 June 2023
Received in revised form 30 September 2023
Accepted 1 November 2023
Available online 11 December 2023

Keywords:
Deep Learning; GAN; Segmentation

ABSTRACT

Lung cancer stands as the foremost contributor to cancer-related fatalities on a global scale, holding the highest mortality rates for both males and females. Tobacco consumption is responsible for approximately 85% of all instances of lung cancer, making it the primary contributor to the disease. Frequently, cancer of the lungs is diagnosed in its advanced stages, limiting treatment options. The screening of individuals at a high risk of developing the disease has the potential to facilitate early detection and significantly enhance survival rates. Regrettably, the statistics remain disheartening because the vast majority of cases are not identified until later stages. Over the past thirty years, medical professionals have explored various approaches to screen individuals deemed to be at a heightened risk of increasing cancer in the lungs. Apart from the fact that CT screening has been shown to lower mortality, there are still issues that lead to unclear diagnoses, needless procedures, financial expenditures, and more. Deep Learning methodologies and tools for data analysis and computing are now widely available. These technical advances will be referenced and utilized in the research to construct prediction models to anticipate the presence of lung cancer in a patient at an early stage. The pictures are pre-processed with the Wiener filter in this proposed method. Following that, pre-processed CT images are segmented with GAN, which is trained using SSSOA (salp shuffled shepherd optimization method), and the segmented output pictures are classified using CNN model VGG16. The F1-score, accuracy, precision, and sensitivity of the various models are compared and evaluated. The VGG16 model predicts lung cancer with great accuracy (97%).

* *Corresponding author.*
E-mail address: rsv.priyan@gmail.com

https://doi.org/10.37934/araset.35.1.4561

1. Introduction

Lung cancer, along with several other cancers, is the leading cause of mortality globally, according to the World Health Organization (WHO). Lung cancer stands as the primary global contributor to mortality, alongside various other forms of cancer. The leading cause of death worldwide has been identified as lung cancer. As per GLOBOCAN data, there is a projected rise of 57.5 percent in new cancer cases in India by 2040, with an estimated 2.08 million new cases, compared to the numbers from 2020. According to the International Agency for Research on Cancer's (IARC) GLOBOCAN 2020 data on cancer incidence and death rates, lung cancer continued to be the leading cause of cancer-related mortality in 2020, accounting for an estimated 1.8 million deaths and 18% of all cancer-related mortalities.

Lung cancer has experienced around 1.8 million fresh diagnoses each year, making up 13 percent of all reported cancer cases, while it has resulted in the loss of 1.6 million lives, constituting 19.4 percent of all cancer-related fatalities. The development of a tumor occurs when abnormal cells persistently multiply and expand. About 85% of male and 75% of female lung cancer cases are linked to cigarette smoke. Lung cancer carries a high mortality rate of 19.4%, particularly impacting low-income countries, making it one of the deadliest diseases.

In Manipur (including Imphal West district) and Tripura, which is in the NER (Northeast region), lung cancer in men was most common. In other regions of the nation, including Delhi, Kollam, Thiruvananthapuram, Bengaluru, Chennai, Kolkata, Nagpur, and Mumbai, the lung scored first among male cancer sites. Only Aizawl reported that the lung was the number one site for females. Nearly two-thirds of tobacco-related malignancies among men in the NER and women in the northern regions of the nation were lung cancers[1].

Symptoms are classified based on the tumor's location and size. In the early stages, it can be challenging to identify since it may not induce pain or any noticeable signs in some instances. People who receive a lung cancer diagnosis can encounter signs such as an ongoing cough, chest pain, breathing difficulties, wheezing, hemoptysis (coughing up blood), Pancoast syndrome (shoulder pain), vocal cord paralysis leading to hoarseness, weight loss, weakness, fatigue, and, in certain instances, passive smoking, which arises from exposure to tobacco smoke. Another contributing factor to lung cancer is genetics. Furthermore, factors such as vehicular pollution, industrial emissions, and exposure to harmful gases like Radon are significant contributors to lung cancer-related deaths, ranking as the second leading cause [21].

One significant contributor to this high mortality rate is the delayed identification of cases. Consequently, a decrease in mortality can be achieved by detecting and addressing cases in their early stages. If the disease is identified early, it can be managed or even completely cured, although this remains a significant challenge for the field of medical science. Late diagnosis, on the other hand, results in the disease reaching advanced, incurable stages, potentially thwarting any successful treatment efforts. Given this context, it's imperative to recognize the importance of technological advancements in automating cancer diagnosis.

It's critical to investigate AI's potential applications for automating the detection of lung cancer as this technology develops, especially in the fields of machine learning and deep learning. Given the growing quantity of accessible training samples and the anticipation of further expansion, supervised learning appears to be an appropriate solution for the automatic discovery of lung cancer. In this setting, an array of methods rooted in image manipulation, machine learning, deep learning, and the field of artificial intelligence (AI) have been created.

Deep learning models heavily depend on the extensive utilization of available data. The most significant challenge when it comes to training a deep learning model is the data acquisition phase.

This challenge becomes even more pronounced when dealing with medical diagnosis models due to the limited availability of internet-based data. Medical information is safeguarded to maintain confidentiality, driven by privacy regulations and apprehensions about the possible misuse of data. Therefore, it is crucial to acquire the training dataset from a reputable and established source as it has a substantial impact on the overall accuracy and correctness of the model. Furthermore, the images collected for training must exhibit a high level of quality to ensure that the model adequately captures all the relevant image features. In this paper, a novel approach is introduced, which utilizes a Generative Adversarial Network (GAN) to segment lung cancer and employs a Convolutional Neural Network (CNN) model using VGG 16 architecture for the detection of lung cancer in the early stage.

The objective of the suggested model is to generate test outcomes that closely resemble actual cases diagnosed by experienced radiologists. As a result, it can function as a valuable instrument to assist healthcare practitioners in their diagnostic decision-making process. Section II conducts a review of existing methods for lung cancer prediction found in the literature. Section III outlines the research methodology employed for a systematic review. Section IV presents the obtained results and engages in a discussion. Lastly, in Section V, the paper concludes and offers insights into potential avenues for future research.

2. Literature Survey

A system was developed to classify and identify nodules in computed tomography (CT) scans. The CT images necessitated the utilization of a computer-assisted detection approach to distinguish between benign and malignant lung nodules, with a strong emphasis on achieving the highest precision to prevent diagnostic delays. Leveraging deep learning technologies for lung nodule categorization consistently produces superior results when it comes to early detection compared to traditional methods.

Rajan *et al*., [12] introduced a kernel-driven method for segmenting lung nodules, joining Binary Firefly Clustering (BFC) with a kernel function in their study. Nonetheless, the challenges stemming from variations in nodule appearance and the heterogeneous nature of lung nodules led to difficulties in achieving accuracy.

Cao *et al*., [7] developed a segmentation of lung nodule approach employing the Dual-branch Residual Network (DB-ResNet). Their method incorporated the Central Intensity Pooling Layer (CIP) to recover intensity features associated with the central region, while Convolutional Neural Networks (CNN) were utilized to extract convolutional features. This approach effectively segmented small nodules as well as nodules located near the pleural region. Nevertheless, it did not achieve full automation of the segmentation process.

Sharma *et al*., [8] introduced a concept called the Synthetic Minority Over-Sampling Technique (SMOTE) for the segmentation of lung nodules. They utilized a blend of SVM (Support Vector Machine) and k-NN (k-nearest Neighbors) classifiers, employing a straightforward sum rule for making predictions. The predictions relied on the cumulative scores from the classifiers. However, because of the imbalanced nature of the dataset, this method struggled to accurately identify the minority class.

In the study by Singadkar *et al*., [6], they utilized a Deconvolutional Residual Network (DDRN) in their strategy for the segmentation of lung nodules. DDRN was employed to capture diverse nodule variations and recover spatial information lost during the pooling operation.

While this method significantly decreased interpretation time, there remains a requirement to enhance the precision of nodule segmentation for all types of pulmonary nodules.

In 2006, Cao *et al*., [7] research team introduced a Two-Stage Convolutional Neural Networks (TSCNN) model specifically designed for lung nodule segmentation. In this case, they utilized the U-Net segmentation network to identify the lung nodules. To reduce false positives, the TSCNN architecture, characterized by dual pooling structures, was employed. This approach effectively segmented smaller local regions. Nevertheless, it's important to highlight that this technique constrained the detection process to a two-dimensional, layer-by-layer approach when segmenting lung nodules.

Mascarenhas and Agarwal [5] illustrated that the VGG16 model outperforms the VGG19 and Resnet50 models in image classification. Notably, it's crucial to mention that this research did not concentrate on a specific image category, despite the unique characteristics present in each type of image. Ismail and Sivothy conducted a comprehensive investigation into breast cancer detection utilizing deep learning methods. They noted that the VGG16 model demonstrated superior performance compared to Resnet50, achieving accuracy rates of 94% and 91.7%, respectively. It's noteworthy that this study specifically focused on unenhanced medical images.

In a notable investigation conducted by Khan *et al*., they explored brain tumor classification utilizing a Convolutional Neural Network (CNN). They created a CNN-based model and carried out a comparative examination against Resnet50, INCEPTION-V3, and VGG16 models. The outcomes of this study revealed accuracy rates of 100%, 89%, 75%, and 96% for the different models, respectively. Munanday *et al*., suggested that this research for detecting facial expressions involves the integration of transfer learning techniques that incorporate convolutional neural networks (CNNs), along with a parameterization approach that minimizes the number of parameters. The FER-2013, JAFFE, and CK+ datasets were jointly used to train the CNN architecture for real-time detection, which broadened the range of emotional expressions that may be recognized. While John and Sujata [23] aimed to detect and locate meningioma brain tumors using the proposed deep learning structure and segmentation algorithm. This proposed MDSS is designed with preprocessing of meningioma and healthy brain MRI images, feature computations, and feature classification through the proposed MENCNN classifier and Meningioma Segmentation Algorithm. The noises in both meningioma and healthy brain images are removed using a Mean Adaptive Filter (MAF) and the meningioma features are computed from the noise-removed image.

Lakshmanaprabu *et al*., [24] asserted that the use of deep learning models had yielded the most effective results in identifying lung cancer from CT data, achieving an impressive accuracy rate of 96.3%, as indicated. In another related study by Lee *et al*.,[25], they investigated the application of deep learning models in chest radiography, particularly in the identification of lung cancer using CT images. The study emphasized the potential of these models to enhance precision and efficacy within clinical contexts.

Abunajm *et al*., [26] introduced a convolutional neural network (CNN) model that aimed to predict and spot lung cancer at an early stage through the analysis of CT scan images. This model demonstrated the ability to distinguish between benign, malignant, and normal cases. The early detection of lung cancer was crucial for enhancing survival rates and enabling timely treatment interventions. The model effectively minimized false positives and attained a remarkable accuracy level of 99.45%.

Based on the literature review, various methodologies have been utilized by different researchers to segment and categorize lung nodules, all with the objective of early-stage lung cancer detection. The analysis highlights that, among these methods, utilizing VGG-16 with deep learning features has proven to be highly effective in categorizing malignant images.

3. Methodology

The proposed method consists of three modules; pre-processing, lung region segmentation, and classification. Figure 1 illustrates the workflow of this study.

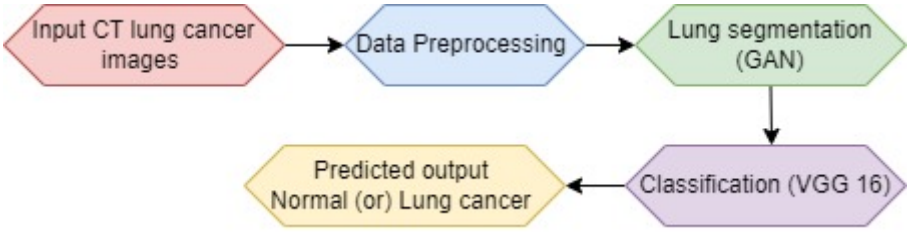

Fig. 1. Pipeline of workflow

3.1 Dataset

In our research, we employed the LIDC/IDRI dataset for the dual purposes of lung cancer segmentation and classification. This dataset was established through a collaboration between the Lung Image Database Consortium (LIDC) and the Image Database Resource Initiative (IDRI), rendering it a valuable and openly accessible resource for the medical imaging research community. Within the LIDC/IDRI dataset, thoracic CT imaging data, complemented by annotations provided by certified radiologists, are made readily available to the public at no cost. This dataset comprises a total of 1018 thoracic CT scans, and each scan is accompanied by XML-based annotations. It was created to facilitate the advancement of computer-aided detection (CAD) methods for various tasks, including lung nodule detection, classification, and quantitative assessment within the field of medical imaging.

3.2 Pre-processing

In image-based classification systems, a pre-processing module is utilized to boost system performance by removing unwanted information, background attributes, and noise. To mitigate Gaussian noise, a Wiener filter is utilized within the pre-processing module.

$$\hat{I}(a,b) = \frac{\sigma_{ab}^2}{\sigma_{ab}^2+\sigma_{nn}^2}\left(S(a,b) - \bar{I}(a,b) + \bar{I}(a,b)\right) \quad (1)$$

Here, σ_{ab}^2 represents the level of variation or dispersion and $\bar{I}$ Stands for the average of the signal without any noise. The description of the signal observation model (s) is as follows:

$$s(a,b) = I(a,b) + e(a,b) \quad (2)$$

In this context, e(a,b) represents the introduced noise with a variance of σ_{nn}^2. Figure 2 illustrates the Lung Cancer CT image input and Weiner Filtered image

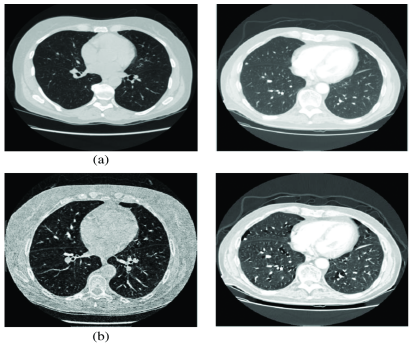

Fig. 2. (a) lung cancer CT image input (b) wiener filtered image

3.3 Lung nodule segmentation using GAN

The schematic representation of SSSOA-based GAN lung region segmentation is illustrated in Figure 3.

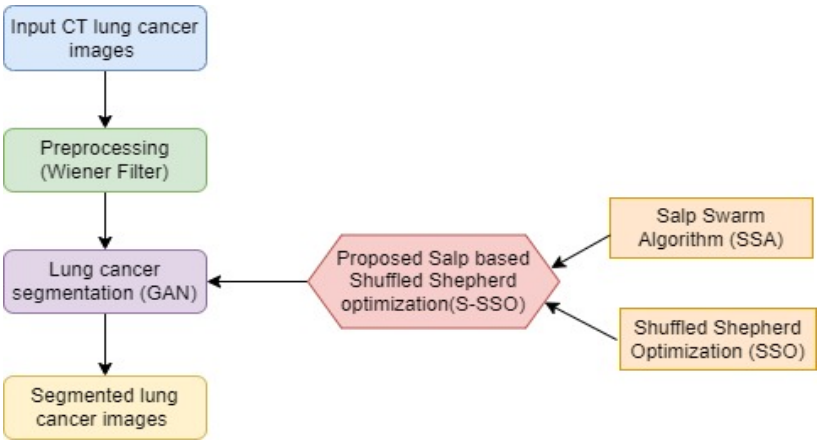

Fig. 3. Schematic representation of SSSOA-based GAN lung region segmentation

A Generative Adversarial Network (GAN) is a deep learning architecture composed of two neural networks, specifically a generator and a discriminator, engaged in competitive training to generate realistic data samples. GANs have found utility across various domains, including their application in tasks like medical image segmentation, such as the segmentation of lung cancer in medical imaging. Figure 4 displays the architecture diagram of the proposed GAN model.

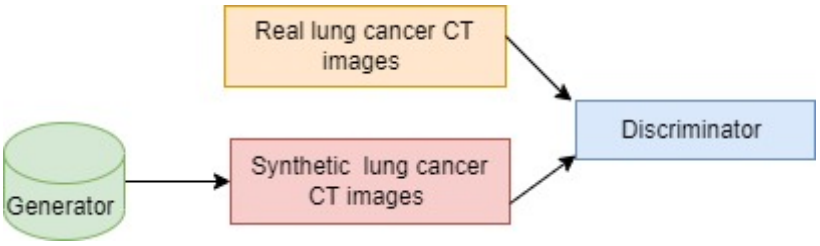

Fig. 4. Proposed GAN architecture

In the context of lung cancer segmentation in medical imaging, GANs can be employed to autonomously detect and delineate the precise locations of lung cancer. This involves tasks like

identifying and outlining these areas within CT scans and making use of Generative Adversarial Networks (GANs). Here's how the GAN framework can help in lung cancer segmentation:

1. Data Preparation: Collect a dataset of medical images (CT scans) that include both healthy lung images and those with lung cancer. These images should be properly annotated with ground truth segmentation masks indicating the regions of lung cancer.

2. Generator Network: The generator's task is to create realistic lung cancer segmentation masks from random noise. It takes random noise as input and generates an output that should resemble a lung cancer segmentation mask. The generator starts with random noise and learns to generate more accurate and realistic masks as it's trained.

3. Discriminator Network: The discriminator functions as a binary classifier to differentiate between authentic segmentation masks sourced from the dataset and counterfeit masks produced by the generator. It receives both real and generated masks as input and produces a probability indicating whether the input mask is genuine or a forgery.

   In recent studies concerning the segmentation of lung cancer images, researchers have placed a significant emphasis on achieving high levels of accuracy. Lung cancer ranks among the most prevalent and fatal forms of cancer. The precise segmentation of lung nodules is a critical step in the accurate diagnosis of lung cancer. Although numerous technologies have surfaced for lung cancer detection, attaining exceptional and accurate detection remains a notable challenge. The difficulties encountered during the lung nodule segmentation process include,

   - The conventional threshold technique generates similar grayscale results for bronchial and tracheal regions but is insufficient for accurate segmentation.
   - The segmentation of lung nodules in CT images is influenced by factors such as their visual attributes, irregular shapes, and the surroundings in which they appear.
   - The segmentation process is influenced by the resemblance in visual features among different lung nodules.
   - Identifying lung nodules reliably poses a formidable challenge due to the intricate surroundings and the diverse nature of lung nodules.
   - The challenges encountered in the current segmentation of lung nodule methods aid as a driving force for further research.

A novel approach named "SSSOA-based GAN" is introduced for the specific purpose of lung nodule segmentation. The primary objective of this study is to develop a lung nodule segmentation technique utilizing the SSSOA-based GAN. The initial stage involves pre-processing, where Wiener filtering is applied to eliminate artifacts present in the original CT image. Subsequently, the pre-processed lung nodule image undergoes segmentation using the SSSOA model, which is trained through a GAN-based approach.

The SSSOA (Salp Swarm Optimization Algorithm) is formulated by combining SSA (Salp Swarm Algorithm) and SSOA (Salp Swarm Optimization Algorithm). SSA is inspired by the collective behavior of Salps as they navigate and forage in the ocean. Within Salp populations, two distinct categories emerge follower Salps and leader Salps. Among these, the leader Salp occupies the foremost position in the chain, while the remaining Salps act as followers. Both groups work collaboratively to track and pursue food sources, with the follower Salps trailing behind the leader Salp.

The SSA exhibits impressive competence in tackling both single-objective and multi-objective problems. In contrast, SSOA takes inspiration from the herding practices of shepherds, leveraging

animal instincts to identify the most optimal pasture choice, thereby requiring fewer evaluations for selection. The seamless integration of SSA and SSOA is employed to determine the optimal solution.

In this research, the straightforward yet highly effective GAN is employed for the segmentation of lung nodules and image synthesis challenges. The GAN model, consisting of a generator, discriminator, and semantic segmentation model, enhances the robustness and accuracy of the segmentation model. To train GAN for nodule segmentation, a hybrid approach known as SSSOA is devised, which combines the methodologies of SSA and SSOA.

Top of FormTraining Process:

- Initialization: The generator and discriminator are initially initialized with random weights.
- Training Alternation: The training process comprises two primary phases, namely the training of the generator and the training of the discriminator.
- Generator Training: The generator is trained to produce segmentation masks that are sufficiently convincing to deceive the discriminator. The generator's loss is determined by how well the generated masks mislead the discriminator.
- Discriminator Training: The discriminator is trained to recognize genuine and produced masks. Its loss is measured based on its capacity to distinguish between real and fake masks.
- Adversarial Loss: The main goal is to reduce the discriminator's capacity to differentiate the generator. This is accomplished using the adversarial loss, in which the generator seeks to maximize the likelihood that the discriminator classifies its output as real.

Within the segmentation model's generator network G, the image is generated, relying on the prediction layer. The discriminator network D plays the role of an overseer in the loss function for G. G has been trained to generate the synthetic image I^S, which corresponds to the original image. Both the original image and the synthetic image I^S are distinguished by D. The specific loss function is articulated within the context of GAN as follows:

$$\text{Loss function} = \min_{\partial g}\min_{\partial d} \log\big(\mathrm{D(I)}\big) + \log\left(1 - (\mathrm{G(I^S)})\right) \tag{3}$$

To circumvent problems stemming from inadequate initialization during the training phase, it's advisable to incorporate a Batch Normalization layer into both the generator and discriminator. In this context, we employ the symbol β to denote a mini-batch containing n examples from the entire training set. Subsequently, the mean and variance of this mini-batch can be computed. Then, we proceed to normalize the data within the micro-batch represented by y_i. Lastly, a transformation step is applied to adjust and reposition the output.

$$\mu\beta \leftarrow \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (\text{mean of mini-batch}) \tag{4}$$

$$\sigma_\beta^2 \leftarrow \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu\beta)^2 \qquad (\text{variance of mini-batch}) \tag{5}$$

$$\hat{y}_i \leftarrow \frac{y_i - \mu\beta}{\sqrt{\sigma_\beta^2 + \epsilon}} \qquad (\text{normalize}) \tag{6}$$

$$Z_i \leftarrow \gamma\hat{y}_i + \beta \equiv CM_\gamma, \beta^{(y_i)} \qquad (\text{scale and shift}) \tag{7}$$

Convergence: As training advances, the generator improves at generating more realistic segmentation masks, while the discriminator improves at discriminating between genuine and

synthetic masks. Ideally, this process leads to a balance where the generator produces masks that are indistinguishable from real masks.
Testing and Inference: After training, the generator can be used for inference. Given a new CT scan image, the generator can produce a corresponding segmentation mask that highlights the suspected lung cancer regions.

3.4 Lung cancer Classification:

The core structure of the CNN architecture encompasses the following layers: Convolution Layer: Convolution layers consist of kernels, Stride, and Padding. Pooling Layer: This layer serves to decrease the computation parameters by reducing the image's dimensions. Fully Connected Layer: Responsible for assigning labels to the images based on the information from the previous two layers. This layer employs the softmax function to ascertain probabilities within the range of 0 to 1.

In our research, we initially employed a customized version of the convolutional neural network model named VGG16. To mitigate potential overfitting issues arising from our relatively small image dataset, we selectively froze certain layers. VGG-16 is renowned for its capabilities in object identification and classification, with the ability to classify 1000 images across 1000 categories. This widely used image classification technique is particularly suitable for transfer learning applications.

Advantages of using the VGG16 model

The VGG-16 model is characterized by its substantial scale, boasting an impressive 138 million parameters. It constructs deep neural networks through the layering of multiple convolutional layers, enhancing its capacity to acquire knowledge about concealed features. When it comes to picture identification and classification, the VGG-16 architecture has a lot of advantages. The VGG-16 architecture is simple to understand and apply due to the use of small 3x3 filters and a simplistic topology. VGG-16 requires fewer parameters than other deep learning architectures, which decreases the computational resources required for training and inference, resulting in lower computational complexity.

3.4.1 VGG-16 Network Model

In 2014, Karen Simonyan and Andrew Zisserman presented the VGG-16 model, which is characterized by a cumulative count of 16 convolutional layers. This model, named after the Visual Geometry Group at Oxford, is characterized by its simplicity and depth compared to AlexNet. In each layer of the network, 3X3 filters were employed with stride and padding sizes set to 1, along with a maximum pooling size of 2. The VGG-16 design consists of a total of sixteen layers, among which there are thirteen convolutional layers intertwined with ReLu layers. It also includes five max-pooling layers and concludes with three fully connected layers, ultimately leading to softmax layers. The VGG 16 Architecture is displayed in Figure 5.

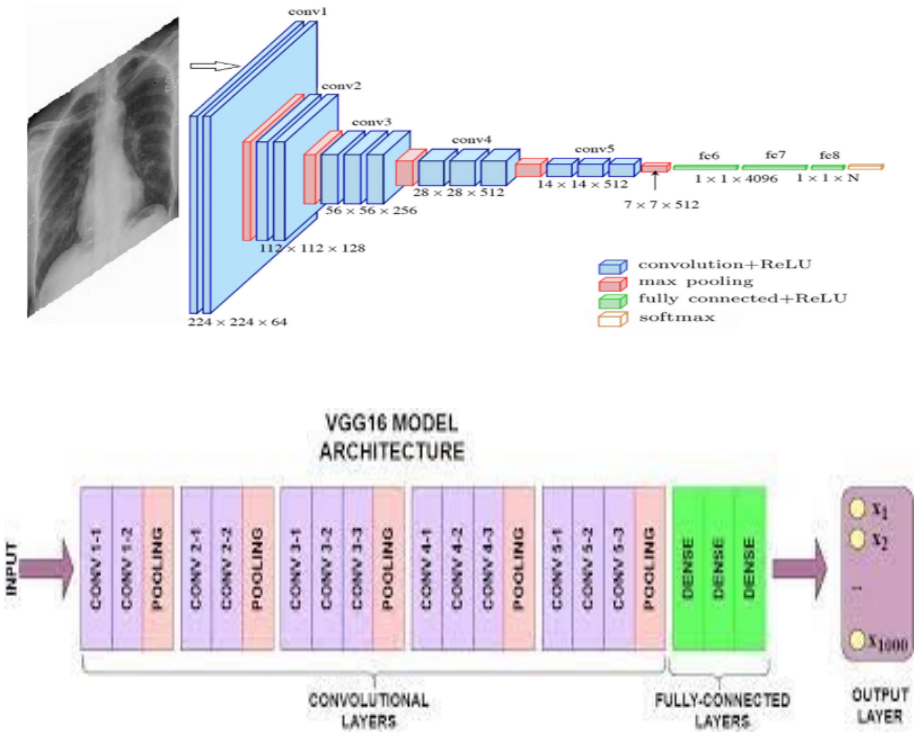

Fig. 5. VGG 16 Architecture

A. Convolutional Layer

This layer serves as the starting point for extracting diverse features from the input images. In this layer, a convolution operation occurs between the input image and a filter of size MxM. The filter moves across the input image, computing the dot product with portions of the input image that match the dimensions of the filter (MxM).

The result of this operation is termed the Feature map, which contains data regarding image characteristics like edges and corners. The following layers then take in this feature map, where they acquire different distinctive features from the input image. The convolution operation preserves the spatial relationships among pixels, a crucial aspect of the convolutional layers in CNNs.

In the cov1 layer, the input consists of a standardized 224 x 224 RGB image. This lung-segmented image undergoes a sequence of convolutional (conv.) layers, using filters of a minimal receptive field size of 3x3, capturing fundamental aspects such as horizontal, vertical, and central information. Furthermore, one of the configurations employs 1x1 convolution filters, acting as a linear transformation of the input channels, followed by a non-linear operation.

The convolutional stride is configured to be one pixel, and the spatial padding of the input to the convolutional layer is adapted to maintain spatial resolution following convolution. This typically entails applying a 1-pixel padding for 3x3 convolutional layers.

B. Pooling Layer

Typically, after a Convolutional Layer, a Pooling Layer comes next. The primary objective of this layer is to reduce the dimensions of the convolved feature map, thus reducing computational requirements. It accomplishes this by reducing inter-layer connections and operating independently on each feature map. Various Pooling techniques are at your disposal, depending on your chosen approach. In essence, these techniques consolidate the features generated by a convolutional layer.

Max Pooling selects the highest value within a feature map, while Average Pooling computes the mean of elements within a specified region of the image. Sum Pooling, on the other hand, calculates the total sum of elements in the designated area.

The Pooling Layer typically acts as a bridge connecting the Convolutional Layer and the Fully Connected (FC) Layer. This arrangement enables the CNN model to distill the features learned by the convolutional layer and facilitates independent feature recognition within the network. Furthermore, it plays a role in lessening computational demands within the network.

In certain instances, five max-pooling layers are employed after particular convolutional layers to carry out spatial pooling, although not all convolutional layers are succeeded by max-pooling. Max-pooling is conducted using a stride of 2 across a 22-pixel window.

C. Fully Connected Layer

The Fully Connected (FC) layer, which includes neurons along with associated weights and biases, plays a crucial role in establishing connections between neurons across different layers. These FC layers are usually situated towards the latter part of a CNN architecture, often positioned just before the output layer. In this process, the input images from the preceding layers are flattened and conveyed to the FC layer.

The flattened vector then proceeds through a series of additional FC layers, where mathematical operations are commonly performed. This signifies the initiation of the classification process. The reason for using two connected layers is that the performance of two fully connected layers exceeds that of a single layer. These CNN layers decrease the need for human supervision.

Following a series of convolutional layers, each with different depths in various architectures, three fully connected (FC) layers are incorporated. The first two of these layers consist of 4096 channels each, while the third is specifically designed for 1000-way classification, containing 1000 channels, each representing a unique class.

D. Dropout

Overfitting is a common issue during training when all features are directly linked to the FC layer. Overfitting occurs when a model becomes excessively skilled at the training data, which can harm its performance when applied to new, unseen data.

To combat this problem, a dropout layer is utilized, which entails randomly deactivating a subset of neurons within the neural network during the training phase, essentially creating a smaller model. For example, when using a dropout rate of 0.3, it implies that 30% of the nodes within the neural network are randomly deactivated.

Dropout contributes to improving the performance of a machine learning model by mitigating overfitting and simplifying the network. It achieves this by selectively deactivating neurons within the neural networks during the training phase.

E. Activation Functions

The activation function holds a crucial role within the CNN model, as it is responsible for capturing and approximating complex and continuous relationships among network variables. Essentially, it determines which information within the model should be passed forward and what should not, introducing a layer of nonlinearity to the network.

Several well-known activation functions exist, such as Softmax, ReLU, Sigmoid functions, and tanH. Each of these activation functions serves a specific purpose. For example, the sigmoid and softmax functions are favored in binary classification CNN models, while softmax is commonly used in multi-class classification scenarios.

In a CNN model, activation functions act as gatekeepers, making decisions about whether a neuron should be activated or not. They evaluate the importance of input data and decide whether it should undergo further mathematical processing. Towards the end of the network, the final layer typically takes the form of a softmax layer. Regardless of the network's configuration, the fully connected layers remain constant. Ultimately, the probability of the presence of lung cancer is calculated by a fully connected layer that includes a softmax layer.

The Model (VGG-16) Procedure:

1. Gather a dataset of Lung Cancer images, encompassing both afflicted and non-afflicted individuals, from the existing LIDC dataset.
2. Pre-process all segmented lung cancer images using the VGG-16 model.
3. Attribute class labels, distinguishing between benign and malignant cases, to the images.
4. Organize the images based on their class labels into training and testing datasets.
5. Utilize 80% of the training images to train the VGG-16 Model.
6. Assess the model's performance by subjecting it to the remaining 20% of testing images.
7. Evaluate various performance metrics.
8. Validate the efficiency and performance of the suggested model.

4. Result and Discussion

Confusion Matrix: A confusion matrix is a structured table employed to evaluate the performance of a classification model. It offers a succinct summary of the model's predictions concerning the actual values of the target variable. The confusion matrix is composed of four essential elements:

1. True positive (TP): Cases in which the model accurately forecasted the positive class
2. True Negatives (TN): Situations in which the model made accurate predictions for the negative class.
3. False Positives (FP): Occurrences where the model projected a positive class outcome despite the true class being negative.
4. False Negatives (FN): Occurrences where the model forecasted a negative class outcome even though the true class was positive.

The confusion matrix provides the basis for calculating several performance metrics, including:

1. Accuracy
2. Precision
3. Sensitivity (also known as Recall)
4. F1 Score

The lung cancer prediction Performance evaluation of various models is provided in Table 1 and the Accuracy, Precision, Sensitivity, and F1-Score performance comparison of various models are illustrated in Figures 6, 7, 8, and 9.

Table 1
Performance evaluation of various models for lung cancer prediction

Models	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-score (%)
ResNet50	93.91	95.74	81.82	96.26
ResNet18	93.04	96.81	84.21	95.79
GoogLeNet	92.54	94.01	80.11	95.25
ImageNet	91.02	92.01	91.05	91.02
VGG16	96.85	97.47	93.21	97.80

4.1 Accuracy

The ratio of accurately anticipated observations with all observations. When the datasets exhibit symmetric false positives and false negatives, it is a useful statistic.
Eq(8) mathematically represents the same thing.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

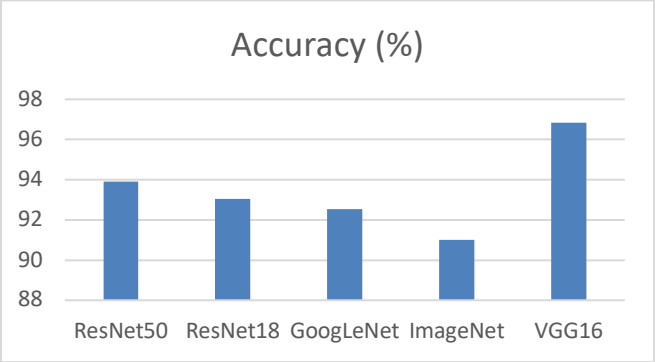

Fig. 6. Accuracy of various models

Figure 5 indicates that the VGG-16 model has the highest level of accuracy (96.85%) when compared to other models.

4.2 Precision

The percentage of correctly predicted positive observations among all observations in the true class is referred to as Precision. It is mathematically expressed by Eq(9) and is related to the false positive rate.

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

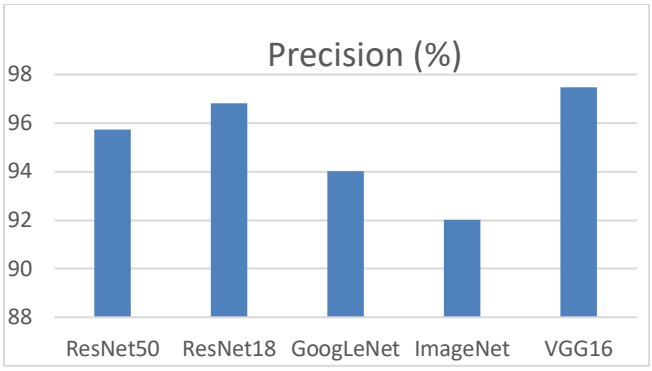

Fig. 7. The precision of various models

Figure 6 indicates that, when compared to other models, VGG16 has a high precision rate (97.47%).

4.3 Sensitivity (or) Recall

The proportion of correctly predicted positive instances relative to the total expected positive instances is known as "Sensitivity" or "Recall." as determined by the formula in Equation (10).

$$Sensitivity = \frac{TP}{FN+TP} \tag{10}$$

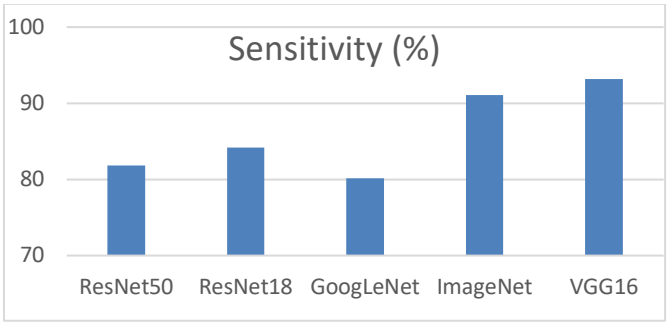

Fig. 8. Sensitivity of various models

This high level of sensitivity is particularly crucial in medical diagnostics, where failing to detect a condition, such as cancer, could lead to severe consequences. The VGG 16 model exhibited a high sensitivity of 93.21% indicating its robust ability to correctly identify cases with lung cancer.

4.4 F1 score

The weighted average of Precision and Recall is typically referred to as the “F1 Score”, especially useful in situations where class imbalances exist or where false positives and false negatives have differing consequences. This is mathematically defined by Equation (11).

$$F1\ Score = \frac{2\times(Precision\times Recall)}{(Precision+Recall)} \tag{11}$$

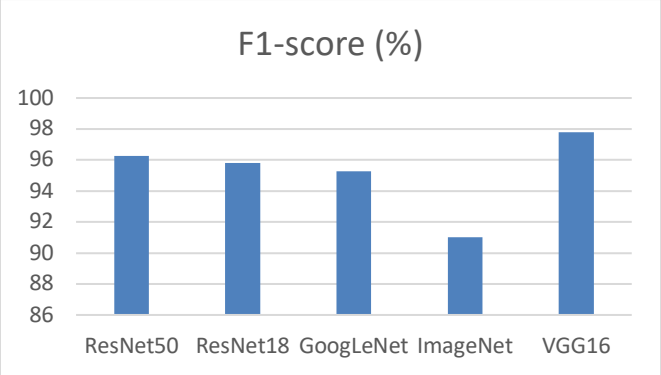

Fig. 9. F1-Score of various models

Figure 8 shows that VGG 16 has a high F1 score (97.80%), indicating that the model excels in correctly identifying both positive and negative instances while simultaneously minimizing the occurrences of false positives and false negatives.

Deep learning has demonstrated its effectiveness in image categorization, with notable models like VGG-16, ResNet50, ResNet18, GoogLeNet, and ImageNet showcasing strong performance. In this research, a dataset obtained from LIDC/IDRI was utilized for both training and testing purposes. This dataset served as input to the network model, designed to identify and classify cancerous (Malignant Images) and non-cancerous (Benign Images).

The VGG-16 model's capacity to detect lung cancer was thoroughly assessed using an extensive range of metrics derived from a confusion matrix. These metrics encompass sensitivity, recall, F1 score, and accuracy. The evaluation of the VGG-16 model culminated in an impressive accuracy of 96.85%, marking the highest level of accuracy achieved in comparison to other models.

5. Conclusion

The VGG-16 model utilized in this study produced good findings, with a 97% accuracy rate in classifying lung cancer. The model's performance was additionally assessed through metrics including sensitivity, recall, F1 score, and accuracy. In comparison to established methods, the VGG-16 model surpassed ResNet50, ResNet18, GoogLeNet, and ImageNet in terms of effectiveness. This highlights the VGG-16 architecture's potential in lung cancer detection, which needs additional investigation. More sophisticated CNN models may provide superior diagnosis accuracy for lung cancer and other medical disorders as the field progresses. As research improves and these models are developed, they have the potential to transform healthcare diagnostic and treatment accuracy.

Acknowledgment

The authors wish to express their thanks to one and all who supported them during this work. The authors received no specific funding for this study. The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Sathishkumar, Krishnan, Meesha Chaturvedi, Priyanka Das, S. Stephen, and Prashant Mathur. "Cancer incidence estimates for 2022 & projection for 2025: result from National Cancer Registry Programme, India." *Indian Journal of Medical Research* (2023).

[2] Wankhade, Shalini, and S. Vigneshwari. "A novel hybrid deep learning method for early detection of lung cancer using neural networks." *Healthcare Analytics* 3 (2023): 100195. https://doi.org/10.1016/j.health.2023.100195

[3] Ismail, Nur Syahmi, and Cheab Sovuthy. "Breast cancer detection based on deep learning technique." In *2019 International UNIMAS STEM 12th engineering conference (EnCon)*, pp. 89-92. IEEE, 2019. https://doi.org/10.1109/EnCon.2019.8861256

[4] Khan, Hassan Ali, Wu Jue, Muhammad Mushtaq, and Muhammad Umer Mushtaq. "Brain tumor classification in MRI image using convolutional neural network." *Mathematical Biosciences and Engineering* (2021).

[5] Mascarenhas, Sheldon, and Mukul Agarwal. "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification." In *2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON)*, vol. 1, pp. 96-99. IEEE, 2021. https://doi.org/10.1109/CENTCON52345.2021.9687944

[6] Singadkar, Ganesh, Abhishek Mahajan, Meenakshi Thakur, and Sanjay Talbar. "Deep deconvolutional residual network based automatic lung nodule segmentation." *Journal of digital imaging* 33 (2020): 678-684. https://doi.org/10.1007/s10278-019-00301-4

[7] Cao, Haichao, Hong Liu, Enmin Song, Guangzhi Ma, Xiangyang Xu, Renchao Jin, Tengying Liu, and Chih-Cheng Hung. "A two-stage convolutional neural networks for lung nodule detection." *IEEE journal of biomedical and health informatics* 24, no. 7 (2020): 2006-2015. https://doi.org/10.1109/JBHI.2019.2963720

[8] Sharma, Srishti, Prasenjeet Fulzele, and Indu Sreedevi. "Hybrid model for lung nodule segmentation based on support vector machine and k-nearest neighbor." In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 170-175. IEEE, 2020. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00034

[9] Cao, Haichao, Hong Liu, Enmin Song, Chih-Cheng Hung, Guangzhi Ma, Xiangyang Xu, Renchao Jin, and Jianguo Lu. "Dual-branch residual network for lung nodule segmentation." *Applied Soft Computing* 86 (2020): 105934. https://doi.org/10.1016/j.asoc.2019.105934

[10] Naik, Amrita, and Damodar Reddy Edla. "Lung nodule classification on computed tomography images using deep learning." *Wireless personal communications* 116 (2021): 655-690. https://doi.org/10.1007/s11277-020-07732-1

[11] Kanipriya, M., C. Hemalatha, N. Sridevi, S. R. SriVidhya, and SL Jany Shabu. "An improved capuchin search algorithm optimized hybrid CNN-LSTM architecture for malignant lung nodule detection." *Biomedical Signal Processing and Control*78 (2022): 103973. https://doi.org/10.1016/j.bspc.2022.103973

[12] Rajan Baby, Yadhu, and Vinod Kumar Ramayyan Sumathy. "Kernel-based Bayesian clustering of computed tomography images for lung nodule segmentation." *IET Image Processing*14, no. 5 (2020): 890-900. https://doi.org/10.1049/iet-ipr.2018.5748

[13] Kamal, Kamal, and EZ-ZAHRAOUY Hamid. "A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images." (2023). https://doi.org/10.21203/rs.3.rs-2863523/v1

[14] Wang, Yunpeng, Lingxiao Zhou, Mingming Wang, Cheng Shao, Lili Shi, Shuyi Yang, Zhiyong Zhang, Mingxiang Feng, Fei Shan, and Lei Liu. "Combination of generative adversarial network and convolutional neural network for automatic subcentimeter pulmonary adenocarcinoma classification." *Quantitative imaging in medicine and surgery* 10, no. 6 (2020): 1249. https://doi.org/10.21037/qims-19-982

[15] Shamrat, FM Javed Mehedi, Sami Azam, Asif Karim, Rakibul Islam, Zarrin Tasnim, Pronab Ghosh, and Friso De Boer. "LungNet22: a fine-tuned model for multiclass classification and prediction of lung disease using X-ray images." *Journal of Personalized Medicine* 12, no. 5 (2022): 680. https://doi.org/10.3390/jpm12050680

[16] Ren, Zeyu, Yudong Zhang, and Shuihua Wang. "A hybrid framework for lung cancer classification." *Electronics* 11, no. 10 (2022): 1614. https://doi.org/10.3390/electronics11101614

[17] Gayathri, Prerepa, Aiswarya Dhavileswarapu, Sufyan Ibrahim, Rahul Paul, and Reena Gupta. "Exploring the Potential of VGG-16 Architecture for Accurate Brain Tumor Detection Using Deep Learning." *Journal of Computers, Mechanical and Management* 2, no. 2 (2023): 23056-23056. https://doi.org/10.57159/gadl.jcmm.2.2.23056

[18] Deepapriya, B. S., Parasuraman Kumar, G. Nandakumar, S. Gnanavel, R. Padmanaban, Anbarasa Kumar Anbarasan, and K. Meena. "Performance evaluation of deep learning techniques for lung cancer prediction." *Soft Computing* 27, no. 13 (2023): 9191-9198. https://doi.org/10.1007/s00500-023-08313-7

[19] Zhang, Yongmei, Bin Dai, Minghui Dong, Hao Chen, and Mengyang Zhou. "A Lung Cancer Detection and Recognition Method Combining Convolutional Neural Network and Morphological Features." In *2022 IEEE 5th International Conference on Computer and Communication Engineering Technology (CCET)*, pp. 145-149. IEEE, 2022. https://doi.org/10.1109/CCET55412.2022.9906329

[20] Ramana, Kadiyala, Madapuri Rudra Kumar, K. Sreenivasulu, Thippa Reddy Gadekallu, Surbhi Bhatia, Parul Agarwal, and Sheikh Mohammad Idrees. "Early prediction of lung cancers using deep saliency capsule and pre-trained deep learning frameworks." *Frontiers in oncology* 12 (2022): 886739. https://doi.org/10.3389/fonc.2022.886739

[21] Ramapraba, P. S., P. Epsiba, K. Umapathy, and E. Sivanantham. "Auxiliary Classifier of Generative Adversarial Network for Lung Cancer Diagnosis." *Intelligent Automation & Soft Computing* 36, no. 2 (2023). https://doi.org/10.32604/iasc.2023.032040

[22] Munanday, Anbananthan Pillai, Norazlianie Sazali, Arjun Asogan, Devarajan Ramasamy, and Ahmad Shahir Jamaludin. "The Implementation of Transfer Learning by Convolution Neural Network (CNN) for Recognizing Facial Emotions." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 32, no. 2 (2023): 255-276. https://doi.org/10.37934/araset.32.2.255276

[23] Anita, John Nisha, and Sujatha Kumaran. "Detection and Segmentation of Meningioma Tumors Using the Proposed MENCNN Model." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 32, no. 2 (2023): 1-13. https://doi.org/10.37934/araset.32.2.113

[24] Lakshmanaprabu, S. K., Sachi Nandan Mohanty, K. Shankar, N. Arunkumar, and Gustavo Ramirez. "Optimal deep learning model for classification of lung cancer on CT images." *Future Generation Computer Systems* 92 (2019): 374-382. https://doi.org/10.1016/j.future.2018.10.009

[25] Lee, Sang Min, Joon Beom Seo, Jihye Yun, Young-Hoon Cho, Jens Vogel-Claussen, Mark L. Schiebler, Warren B. Gefter et al. "Deep learning applications in chest radiography and computed tomography." *Journal of thoracic imaging* 34, no. 2 (2019): 75-85. https://doi.org/10.1097/RTI.0000000000000387

[26] Abunajm, Saleh, Nelly Elsayed, Zag ElSayed, and Murat Ozer. "Deep Learning Approach for Early Stage Lung Cancer Detection." *arXiv preprint arXiv:2302.02456* (2023).