



Squeeze-Inception V3 with Slime Mould algorithm-based CNN features for lung cancer detection

Geethu Lakshmi G^{*}, P. Nagaraj

Department of Computer Science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Srivilliputhur, Tamil Nadu, India



ARTICLE INFO

Keywords:

Lung cancer
CNN features
SqueezeNet
Inception V3
Optimization algorithm

ABSTRACT

Worldwide, lung cancer is considered a highly fatal disease, affecting both women and men; it needs additional consideration while assessing several diseases. Moreover, early lung disease detection is highly significant in raising the survival rate of affected persons. There are numerous approaches to detecting lung cancer, yet it can be highly complicated to place the affected region because of the minimal visibility of the tumor section and error rates. In order to diagnose lung cancer accurately, this work created a technique, named as Squeeze-Inception V3 with Slime Mould Algorithm (SMA)-based Convolutional Neural Network (CNN) features. Here, SMA-CNN is proposed for extracting CNN features from preprocessed images since training CNN using SMA enhances the performance. Subsequently, the features are fed to the classification for detecting lung cancer. The classification into normal/abnormal is done using the proposed Squeeze-Inception V3, which is designed by integrating the SqueezeNet and Inception v3 architectures. Finally, the analysis is performed with accuracy, sensitivity, as well as specificity rates of 95 %, 98 %, and 93 %, and reveals that the performance of proposed technique is superior to comparative models.

1. Introduction

In several countries, lung cancer is considered most dreadful illness, and its death rate is high. The numerous techniques utilized by radiologists to recognize cancer involve CT scans, sputum cytology, X-rays, and other MRI approaches. Tumors are classified into two categories, benign and malignant during the detection procedure. Generally, malignant tumors are proliferating and cancerous and have irregular sizes and shapes [1]. Also, analysis exhibits that the patients with advanced stage have a substantially lower survival rate than those with early-stage disease. Additionally, it has been evaluated that timely analysis of the imaging and scans can be enhanced by using several image processing approaches. Several studies were developed for recognizing early-phase cancers by utilizing image processing techniques. Because of two major issues, hit ratio of manually detecting lung cancer might be minimized. The foremost challenge is human and technical accessibility, as there

might not be adequate radiology resources to fulfil the demand. The next one is that a large amount of false positive scenarios because of first drawback. Thus, higher quality training must be given to radiologists interpreting images. Therefore, the accuracy of recognition and classification for conventional approaches still needs enhancements [2].

The prevention, control, and treatment of respiratory disorders are significant because they represent five out of the top thirty most general causes of death. To assist in fast recovery and to increase the survival rates, early diagnosis is required. Conventionally, CT scans and chest radiographs (X-rays) are utilized for detection of lung disease. Generally, chest X-rays are low-cost, simple to operate, extensively accessible, and more time-effectual than CT. They comprise a large number of information regarding the health of patients [3]. In spite of these benefits, interpreting them remains an issue. Even skilled radiologists have difficulty identifying obscure nodules or differentiating between similar lesions [4]. Moreover, manual screening of lung diseases is labor

Abbreviations: CAD, Computer-Aided Detection; HRCT, High-Resolution Computed Tomography; DL, Deep Learning; ML, Machine Learning; GANs, Generative Adversarial Networks; SMA, Slime mould algorithm; LDA, Linear Discriminant Analysis; MRI, Magnetic Resonance Imaging; DCNN, Deep Convolutional Neural Network; FPSO, Fuzzy Particle Swarm Optimization; E2E, End-to-End; SIFT, Scale Invariant Feature Transform; E-nose, Electronic Nose; LBP, Local Binary Pattern; SGL, Spare Group Lasso; CNNs, convolutional neural networks; LR, linear regression; SVM, Support Vector Machine; CT, Computed Tomography; DT, Decision Tree; ReLU, rectified linear unit; HoG, Histogram of oriented Gradients; TLEV2N, transfer learning-based EfficientNetV_2 network; kNN, k-nearest neighbor; H&E, hematoxylin and eosin; GS-PCA, Graph-Based sparse Principal Component Analysis; FexViT, feature extraction Vision Transformer.

* Corresponding author.

E-mail addresses: geethu010@gmail.com (G. Lakshmi G), nagaraj.p@klu.ac.in (P. Nagaraj).

intensive, time utilizing, as well as undergoes intra and inter-observer differences [5]. In the early phase, identifying lung cancer is the only technique for its cure. Diverse techniques are present to diagnose lung cancer, such as isotope, MRI, CT, and X-ray. CT and X-ray chest radiography are the two renowned anatomic imaging modalities, which are often employed in recognition of diverse lung diseases. Radiologists and physicians use CT scans to locate and diagnose the disease, directly depict the morphologic extents of diseases, evaluate the patterns and severity of the diseases, and evaluate clinical course of disease and response to medicine. The volumetric CT approach has developed spiral scans that reduce scan time and while utilized in thoracic imaging, minimize artefacts caused by limited volume effects, unequal respiratory cycles, and cardiac motion. With the development of CT technology, high-resolution CT examination has occurred as the imaging modality option for the identification and recognition of lung disease. Although HRCT suggests images of lung with increasingly enhanced visual interpretation, anatomic resolution, or assessment of a huge number of CT image slices remains a complicated task [6].

Current developments in DL and ML approaches have shown an important transfer towards CAD systems for lung cancer detection [2]. In DL, the current advancements have exhibited enormous possibilities in automatic recognition and classification of patterns in medical images [7]. Particularly, CNN has attained outstanding performance in detection of diseases, namely lung diseases. Nonetheless, these approaches are a great deal dependent on a huge number of labeled training data or fine-tuning the millions of parameters from pre-trained CNNs. Additionally, number of annotated data accessible in medical imaging field is restricted, particularly for a new disease. Even though the auto annotation platforms are accessible for lung disease, the accuracy and effectiveness of these systems are minimal, and they are still in primary phase of investigation [8]. The aforementioned problems restrict performance of aforesaid supervised learning techniques on new data. In contrast, unsupervised representation learning approaches are equivalent to the performance of supervised approaches by utilizing only unlabelled data. One kind of unsupervised learning approach is the autoencoder in which the encoder uses input data and outputs a compressed illustration of input [9]. Subsequently, decoder utilizes this demonstration to recreate the initial input data. Nevertheless, aforesaid approaches often exhibit inadequate performance in image generation, resulting in frequently blurry and not good quality reconstructed images because of compression. GANs [10], conversely, present an unsupervised approach for generating new image samples without a requirement for supervision, by employing a two-player min–max game and artificial expansion of the training dataset. Currently, an alternative to GANs, known as deep convolutional GANs, shows promise in the realm of image synthesis tasks. They demonstrate the capability of GANs to capture intricate data distributions and learn more effective image representations.

This work proposes a hybrid deep-learning algorithm for lung cancer detection. Here, the images undergo a series of preprocessing steps, such as color space conversion, data augmentation, resizing, and normalization. The resulting pre-processed images are subsequently passed via a feature extraction phase, where a CNN, trained with the SMA, is utilized to extract relevant features. These extracted features are then input into a classification process to identify lung diseases. The classification task distinguishes between normal and abnormal cases, which is accomplished using the Squeeze-Inception V3 model.

The major contributions of the proposed model are as below:

- To extract relevant patterns for lung disease detection, the CNN features optimized by the SMA are designed.
- To classify the image as normal and abnormal, Squeeze-Inception V3 is developed by combining the architectures of SqueezeNet and Inception V3 so that the method utilizes the advantages, like accuracy, high performance and low computational speed, of both models.

- To analyze performance of proposed approach, measures, namely accuracy, specificity, and sensitivity, are employed.

2. Literature review

This section explains a general idea of previous studies and research performed in field of lung cancer detection.

Pooja Yadav *et al.*, [5] designed a deep unsupervised model to classify lung diseases from CT images and chest X-rays. Here, architecture, called Lung-GANs, represented as multi-layer generative adversarial networks, was introduced. The objective behind these networks was to acquire meaningful illustrations of lung disease images through the utilization of unlabeled data exclusively. The lung features acquired by the approach were used to train Stacking Classifier and SVM. The experimentation evaluation states that performance of the approach was better than unsupervised approaches in lung disease classification. Rabbia Mahum and Abdulmalik S. Al-Salman [2] developed an effective lung tumor detector on the basis of RetinaNet, called Lung-RetinaNet. To aggregate several network layers, a multi-scale feature fusion-based phase was developed, concurrently raising semantic information from shallow prediction layer. In addition, a lightweight and dilated approach was used for context phase to integrate the contextual information with each network phase layer to enhance features as well as efficiently localize small tumors.

A. Asuntha and Andy Srinivasan [6] detected cancerous lung nodules from a given input lung image. A DL model was used to recognize the cancerous lung nodule's location. Here, optimal feature extraction approaches, like wavelet transform-based features, SIFT; HoG, LBP, and Zernike Moment were used. Subsequent to extracting texture, geometric, volumetric and intensity features, the FPSO approach was used to choose the best features. At last, these features were classified by utilizing the Deep learning model. Bei Liu *et al.*, [11] developed an E-nose system with a preconcentration subsystem, improved with SGL feature selection approach. Conversely, a preconcentration subsystem integrated with two conformal sensor chambers was developed. A kind of multi-bed sorbent tube was used for an enhanced desorption and adsorption effect on VOCs in the preconcentration subsystem. Conversely, SGL was developed to enhance and extract helpful information from raw data, based on multiple kinds of features.

Mesut Togaçar *et al.*, [12] realized lung cancer detection by utilizing the AlexNet, LeNet, and VGG-16 DL approaches. The experimentations were performed on an open dataset comprised of CT images. Here, CNNs were employed for feature extraction and classification functions. During the training of the approaches, to raise success rate of classification, image augmentation approaches, like cutting, horizontal turning, zooming, and filling, were used in dataset. Owing to exceptional achievement of AlexNet model, features achieved from last fully connected layer of the model were separately applied as the input to LR, DT, SVM, kNN, LDA, and softmax classifiers. Sundaresh Ram *et al.*, [13] developed a basic ML model, named GS-PCA model for the automated identification of malignant growths in histological lung tissue samples stained with H&E. The method consists of four phases, such as PCA binary hashing, cascaded graph-based sparse PCA, blockwise histograms, and SVM classification. Here, graph-based sparse PCA was used for learning filter banks of multiple phases of a convolutional network.

Onur Ozdemir *et al.*, [14] worked on a CAD and identification model for lung cancer screening with low-dose CT scans, which generate significant probability evaluations. The technique was based completely on 3D CNN and obtained superior performance for both malignancy classification tasks and lung nodule detection. Also, the coupling between the analysis and detection models was very important to consider. Using these coupling permits to design of an E2E model, this has better and reliable performance and eradicates the requirement for a nodule detection and false positive minimization phase. Lam Pham *et al.*, [15] presented and explored a robust DL model for auscultation analysis. The technique initiates a front-end feature extraction process, which

Table 1
Advantages and limitations of the model.

Authors	Methods	Advantages	Disadvantages
Pooja Yadav <i>et al.</i> , [10]	GANs	It offers higher sensitivity in lung diseases and better performances than other leading unsupervised techniques.	The robustness of the GANs is seen on many datasets of similar lung diseases, the datasets are not real-time patient samples.
Rabbia Mahum and Abdulmalik S. Al-Salman [2]	RetinaNet	It enhances the capability of the proposed network to extract valuable features.	The ability to identify tiny tumors is reduced due to the limited resolution of the input images
A. Asuntha and Andy Srinivasan [6]	FPSO algorithm	It reduces the computational complexity.	It is difficult to grade the images based on the degree of the malignancy of pulmonary nodules.
Bei Liu <i>et al.</i> , [11]	3D PCA and SVM	Final performance shows a steady improvement.	Sample amounts are relatively small.
Mesut Togaçar <i>et al.</i> , [12]	CNNs	The model is a consistent diagnosis model for lung cancer detection using chest CT images.	The method does not generalize to the design of high-performance CAD systems for other medical imaging tasks.
Sundaresh Ram <i>et al.</i> , [13]	GS-PCA network	This method is relatively very fast to learn a good model in comparison to other methods.	Time complexity results in noise and outlier image patches.
Onur Ozdemir <i>et al.</i> , [14]	CNNs	It can effectively used to evaluate lung CT scans.	Failed to visually analyze the learned feature representation to assess and interpret the CAD results
Lam Pham <i>et al.</i> , [15]	CNN-MoE	Achieved high accuracy	High complexity

converts input sound data into a representation of a spectrogram. After that, a front-end DL network was employed to classify these spectrogram features into categories related to respiratory anomalies or cycles. The experimentation was conducted using the ICBHI database of respiratory sounds, confirming the significance of three major contributions to the field of respiratory sound analysis.

Resham Raj Shivwanshi and Neelamshobha Nirala [16] aimed to develop an advanced model to assess lung nodules by employing the automated approaches with CT images for detecting lung cancer at an earlier phase. The approach uses a fixed-size 3×3 kernel in a CNN for appropriate feature extraction. The nodule detection model was improved by combining a TLEV2N to enhance training performance. The nodules classification was attained by combining the EfficientNet_V2 architecture of CNN for more precise malignant and benign classification. Resham Raj Shivwanshi and Neelamshobha Nirala [17] employed FexViT and Feature selection by Quantum Computing based Quadratic unconstrained binary optimization (QC-FSelQUBO) approach. This approach exhibits enhanced performance while comparing with other conventional approaches. The technique exhibited enhanced performance while comparing with other approaches. Table 1 demonstrates the advantages and disadvantages of the models.

2.1. Challenges

The challenges emphasized in previous section have led to the identification of the following research gaps.

- Even though RetinaNet detects the tumors by extracting valuable features from the images; the capability to recognize small tumors is decreased because of the limited resolution of input images [2].
- Moreover, 3D PCA and SVM [11] suffer from a relatively small sample size.
- Another challenge noted is in GS-PCA, where the time complexity of the method resulted in the presence of noise and outlier image patches.
- Furthermore, the inability to visually analyze the learned feature representation by the CNN [14] delays the assessment and interpretation of CAD outcomes.

3. Proposed methodology

A Deep learning approach is proposed in this work for lung disease recognition employing Chest CT-Scan images. The overall framework for lung disease classification is exhibited in Fig. 1. Initially, the images are pre-processed using Color space conversion, data augmentation, resizing, and normalizing [10]. The pre-processed output is fed to feature extraction, where CNN [3], trained with SMA [18] is employed for extraction of features. Then, features are given to the classification for detecting lung diseases. The classification into normal/abnormal is done using the proposed Squeeze-Inception V3, which is designed by combining SqueezeNet and Inception v3 architectures.

3.1. Pre-processing

The pre-processing is done on lung images to eradicate the noise in an image and to enhance image contrast. The main aim of the pre-processing is smoothing input image. Here, pre-processing is used by using color space conversion, data augmentation, normalizing and resizing. The images in the databases are in a pixel range of [0, 255] and have diverse sizes as well as color modes. Since the original image size varies for each image (for example 282×381 , 335×486 , 357×483) each image is normalized to the 0 to 1 range, transformed into RGB, and resized into 256×256 pixels.

3.1.1. Color space conversion

This involves transforming the color representation from one form to another. Generally, this happens in context of transforming an image, which is indicated in one color space to another color space. Moreover, the main aim of the color space conversion is to create the image in a way that makes it closely resemble the original [19]. Here, it is converted into YCbCr, whereas YCbCr color model is indicated regarding two chrominance modules (Cr and Cb) and one luminance component (Y). In this color space, the light intensity is exhibited as "Y" modules. The intensities of the red as well as blue modules relative to the green modules are shown as "Cr" and "Cb" modules correspondingly.

3.1.2. Data augmentation

The data augmentation model is used to raise a number of images by changing conventional dataset to make an artificial dataset. Here, the three most adequately utilized augmentation approaches, vertical flip, horizontal flip, and 90° rotation are used. The input image is flipped by the horizontal flip augmentation along its vertical (left to right) axis arbitrarily with a given probability. Also, input image is flipped by the vertical flip augmentation along its horizontal (top to bottom) axis arbitrarily with a given probability. Oversampling is performed in an image to balance class distributions in a dataset. Generally, oversampling an image involves increasing its size or resolution to make it larger than its original dimensions. The main aim of augmenting original data is to augment data size and enhance the learning capability of DL approach. Subsequently, to prepare the images for DL models and train model faster. After augmentation, the final preprocessed images are 7007 [20].

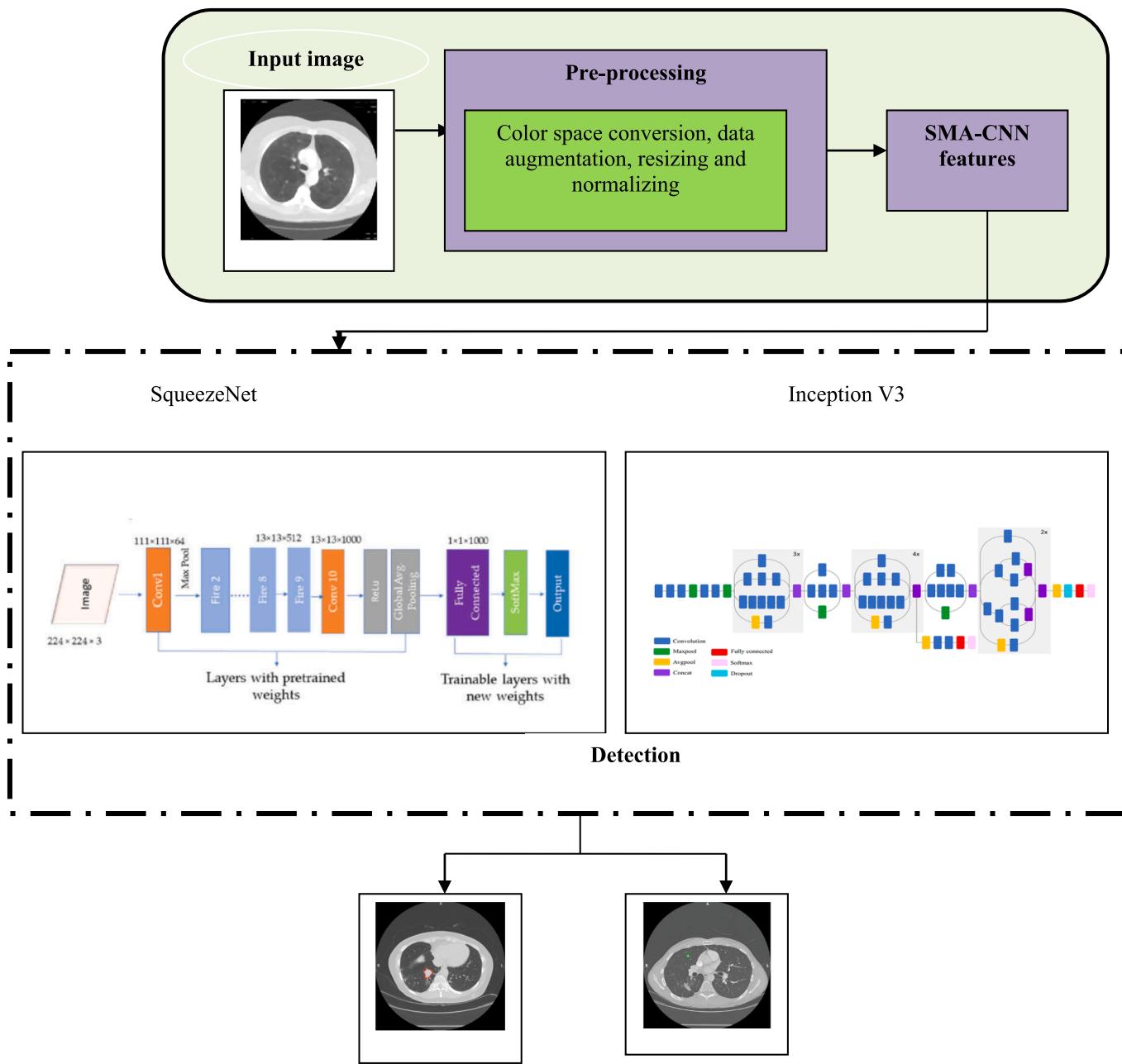


Fig. 1. Architectural diagram of lung disease detection using deep learning model.

3.1.3. Resizing and normalizing

The image from several detection methods has been resized by utilizing nearest neighbor interpolation approach with a specified output size [21]. Generally, resizing and normalizing are frequently carried out in the pre-processing phase, and it involves a process of adjusting the dimension in a standardized way and also that changes pixel intensity values range between 0 and 1.

3.2. SMA-CNN features

This section elaborates on the detailed study of the SMA-based CNN features. Here, features are extracted by utilizing CNN model, which is trained by the SMA approach. The major benefit of CNN features is their capability to learn directly from raw pixel data, which eliminates the need for manual feature engineering. Moreover, training it using the optimization algorithm further improves the extraction performance. Thus, the SMA-CNN feature is used in this work.

3.3. CNN features

CNN is an important feature, which is extracted from the pre-processed image for lung cancer detection. CNN comprises three layers, like pooling layer, the convolution layer, and the fully connected layer. In CNN framework, primary phase is convolutional layer, in which CNN features are extracted by utilizing the pre-processed outcome. The association between pixel value and image attributes is preserved in a convolutional layer [22].

Convolution layer: Basically, convolution layer filters image using a filter/mask. The output is derived from the convolution layer in the form of a matrix.

Activation function: The matrix generated by the convolutional layer is converted into the vector form. Basically, to convert the matrix into vector form, ReLU is exploited as an activation function.

Pooling layer: It is employed to reduce parameter size. Generally, min, max, and average pooling are utilized for feature reduction.

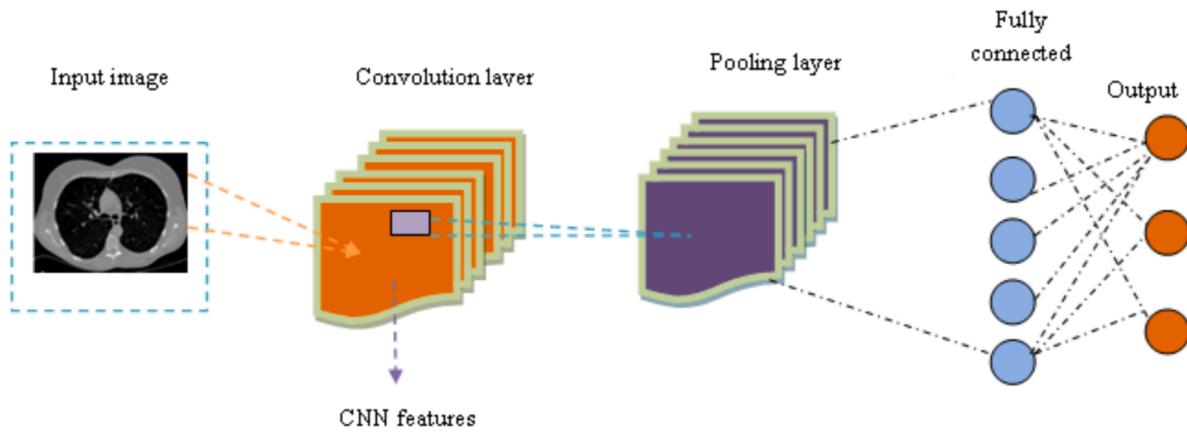


Fig. 2. Architectural model of CNN features.

Dropout: In some cases, the training data may not be suitable for testing procedures. The dropout approach is utilized to resolve overfitting and underfitting problems.

Fully connected layer: The convolutional layer or final pooling that is gentled and positioned in a fully linked layer generates the input.

Softmax: It is an activation function that assists in obtaining FCN output appropriately. The activation function's result from FCN is in the accurate form for a probability distribution. It can be in the form of features either color, shape, or texture forms. As a result, these features are utilized for prediction and detection. Fig. 2 demonstrates the architectural model of feature extraction using CNN.

3.4. Training using SMA

This sub-section describes the algorithm used to tune the CNN, i.e., SMA that is enthused by the natural fluctuation in the slime mold performance [23]. SMA integrates various novel features and unique parameters, which employ adaptive weights to replicate the biological waves. By using its abilities for exploitation and exploration, the SMA presents an ideal path to connect food sources [18]. Slime mould can move towards food on the basis of odor in the air. The below mathematical expression demonstrates how slime mold position is updated at the time of the search procedure,

$$\vec{H}^* = \begin{cases} m \cdot (B_u - B_l) + B_l, & m < y \\ \vec{H}_a(t) + \overrightarrow{\nu k} \cdot (\vec{W} \cdot \vec{H}_M(t) - \vec{H}_Q(t)) & R_n < P_r \\ \overrightarrow{\nu m} \cdot \vec{H}(t) & R_n \geq P_r \end{cases} \quad (1)$$

The formulation for P_r and l represented as follows:

$$P_r = \tanh|f_s(c) - V_f| \quad (2)$$

$$l = \text{arctanh}\left(-\left(\frac{t}{\text{max_}t}\right) + 1\right) \quad (3)$$

where, B_l and B_u signifies lower and upper boundaries of search range, m signifies arbitrary value in $[0,1]$, \vec{H}_a represents current position with highest concentration of odor, t signifies current iteration, $\overrightarrow{\nu k}$ signifies a parameter within the range of $[-l, l]$, $\overrightarrow{\nu m}$ reduces linearly from "1 to 0", \vec{H} signifies location of slime mould, $\vec{H}_M(t)$ and $\vec{H}_Q(t)$ signifies two randomly chosen from slime mould, R_n implies arbitrary value in interval of $[0,1]$. Algorithm 1 represents the pseudo code of SMA model.

Algorithm 1: Pseudo code of SMA model

Initialize the parameters and agent's position
For each agent

(continued)

Algorithm 1: Pseudo code of SMA model

```

Initialize agent location randomly
while
    Compute current position
    if
        Compute movement probabilities on the basis of fitness
        Choose next position
    End if
End while
End for

```

3.4.1. Cancer detection using proposed SqueezeNet-Inception V3

In this section, the cancer detection performed using the Hybrid SqueezeNet-Inception V3 is detailed. The main benefit of the SqueezeNet model relies on its capability to maintain a balance between accuracy and computational resources. Conversely, Inception V3 mainly focused on reducing the computational time by enhancing the preceding Inception framework, which results in higher efficiency and enhanced network convergence. Thus, the SqueezeNet and the Inception V3 are hybridized in this work for lung cancer detection.

SqueezeNet: SqueezeNet is a convolution network that achieves superior performance compared to AlexNet with 50x parameters. Generally, the SqueezeNet [24] consists of several fire modules, and the fire modules comprise an expanded layer and squeeze convolution layer positioned between the convolution layers. The result of the squeeze convolution layer is transmitted to the subsequent expand layer in the fire modules. The SqueezeNet possesses only a few parameters. However, it can resolve the issue of the minimized computational efficiency of the network approach while maintaining an equivalent level of accuracy. In addition, the SqueezeNet begins with an independent convolutional layer, followed by eight fire modules, and finishes with the ultimate convolutional layer. Subsequently, the max-pooling operation is performed by the SqueezeNet, which is tracked in two strides. The structure of fire module is designed to minimize number of approach parameters and maintain accuracy of detection by minimizing the size of convolutional kernel approach, which replaces a fully connected layer with an average pooling layer.

The input images undergo a process of generalization through convolution, which is followed by the application of max-pooling. In the convolution layer, small regions in input volumes are convolved by 3×3 kernels. Additionally, an element-wise activation function is performed that captures the positive part of the input's values. The output tensor scale and the input to fire layers remain consistent. In squeeze phase, a 1×1 filter is used, while the expansion phase utilizes filters of both 1×1 and 3×3 sizes. The squeeze operation minimizes depth of the tensor, whilst the expansion phase raises the depth while maintaining the same

(continued on next column)

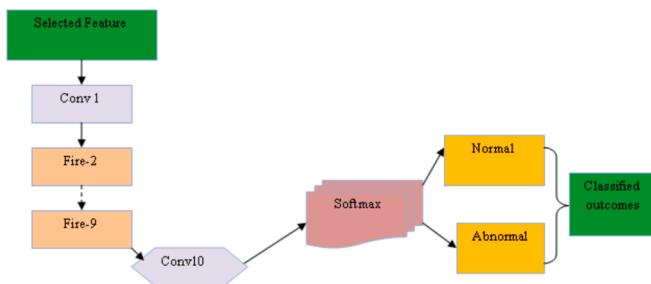


Fig. 3. Architecture model of SqueezeNet.

feature size [25]. Delaying the pooling operation and integrating the pooling layer in final phase of network enhances accuracy of the detection, specifically while number of parameters is restricted. Because pooling layer significantly minimizes the size of the outputs feature map, pooling layer's lag can achieve a higher output feature map, thus, this approach preserves a greater number of features extracted by the convolutional layer. At last, the outputs from the expansion phase are integrated into the depth dimension of the input tensor through a concatenation operation. Fig. 3 demonstrates the architectural model of SqueezeNet.

3.5. Inception V3

InceptionV3 is a DCNN model, which was developed by Google. It is a variant of the original Inception network, designed to attain superior accurateness on image detection tasks. InceptionV3 integrates the idea of Inception phases that utilize multiple parallel convolutional filters of diverse sizes to capture information at several scales. InceptionV3 has been extensively utilized in computer vision applications, namely image segmentation, image recognition, and object detection. The first generation of non-sequential CNNs involved the GoogleNet which took first place in the competition's classification and detection tracks. This Inception V3 network permitted for an increase in both width (the number of units at each level), and depth (the number of levels) without causing the computational processing power [26]. Inception V3 was designed based on the theory that various connections between layers are redundant and ineffectual information because of their correlation. For that reason, it utilizes a sparse CNN, named Inception module with 22 layers in a parallel processing workflow and makes use of various auxiliary classifiers within in-between layers to enhance discrimination ability in lower layers. Unlike conventional CNN model such as VGG and

ALexNet in which either a pooling or a convolutional operation can be utilized at each level, the Inception phase could advantage of both at each layer.

In addition, convolutions (filters) with different sizes are utilized at a similar layer that provides more detailed information and extracts patterns with varying sizes. Significantly, the often referred to bottleneck layer, a 1×1 convolutional layer, is used to minimize both number of parameters and computational complexity. To be more specific, 1×1 convolutional layers were utilized just before a higher kernel convolutional filter (for instance 3×3 and 5×5 convolutional layers) to minimize number of parameters to be ascertained at each level, which is pooling feature process. Additionally, 1×1 convolutional layers create a network deeper and also increase the non-linearity, by employing ReLU subsequent to each 1×1 convolutional layer. An average pooling layer takes the place of the fully connected layers in this network. This considerably minimizes the number of parameters as fully connected layers involve a huge number of parameters. Therefore, this network is capable of learning the deeper representations of features with few parameters in relation to AlexNet when it is much speedier than the VGG. Fig. 4 illustrates the architectural model of Inception V3 model, where R represents reduction block, and S represents stem block. Reduction Blocks are used to modify the dimensions both in terms of width and height of the grid.

The extracted features are passed through both the SqueezeNet and Inception V3 models independently. Subsequently, the output features from both models are then combined using a merge layer. This merging involves operations, like addition, and the main goal of this merging is to combine the strengths of both SqueezeNet and Inception V3 models, potentially capturing a more comprehensive representation of the input data. Finally, the merged features are used for the further task, i.e. lung cancer detection. In Inception V3, Inception phases aspire to balance representational power and computational effectiveness by employing factorized convolutions the Squeeze-Inception V3 network design integrates the advantages of SqueezeNet's parameter efficiency and Inception V3's multi-scale feature extraction abilities. This hybrid model aims to deliver a powerful yet compact DL model appropriate for proficient and effectual lung cancer detection.

4. Result and analysis

The implementation of proposed approach was experimented with in the PYTHON tool as well as "Chest CT-Scan images Dataset [27]" was used for analysis. The system requirements include Windows 10 as the operating system, 8 GB of RAM, 100 GB for storage, a GPU, and a CPU

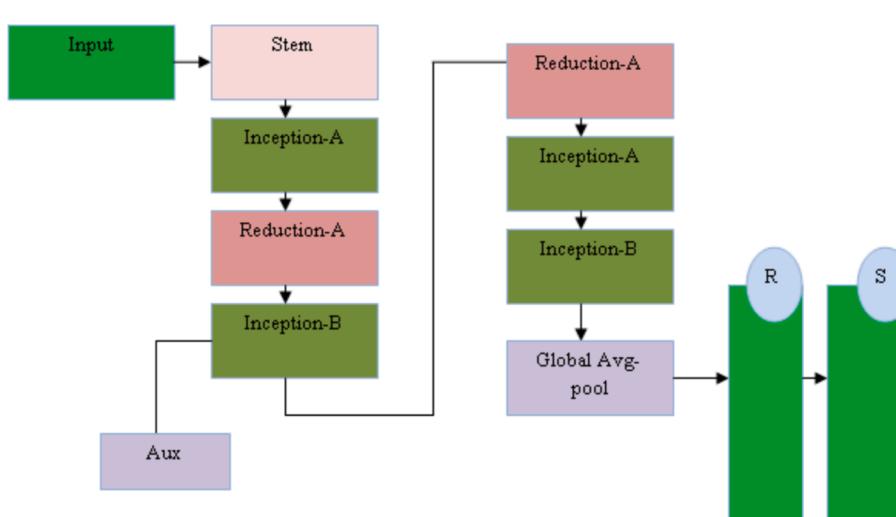
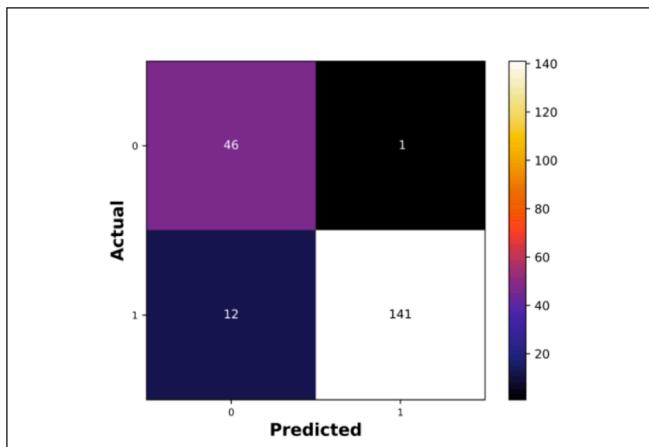


Fig. 4. Architectural model of Inception V3.

**Fig. 5.** Confusion Matrix.

with a minimum speed of 1.7 GHz. In addition, the software requirements consist of Python version 3.7.6 and PyCharm version 2020. Also, analysis of proposed scheme was performed based on evaluation measures, like accuracy, sensitivity, and specificity, and was evaluated with traditional GAN [5], CNN [12], Lung-EffNet [28], and RetinaNet [2].

4.1. Dataset description

The Chest CT-Scan images dataset was obtained from a publicly

accessible source on Kaggle [27]. This lung cancer dataset was meticulously compiled by manually collecting images from various websites, with each label being thoroughly validated. The dataset comprises a total of 1001 CT scan images, categorized into four distinct classes to diagnose lung cancer. These classes include adenocarcinoma, large cell carcinoma, squamous cell carcinoma, and cases categorized as normal.

4.2. Performance measures

4.2.1. Accuracy

Accuracy is a measure that exhibits the percentage of correct classifications and it can be formulated as:

$$\text{Accuracy} = \frac{A + B}{A + B + C + D} \quad (4)$$

Here, A represented as true positive, false positive is exhibited as C , B represents true negative, and false negative is signified as D .

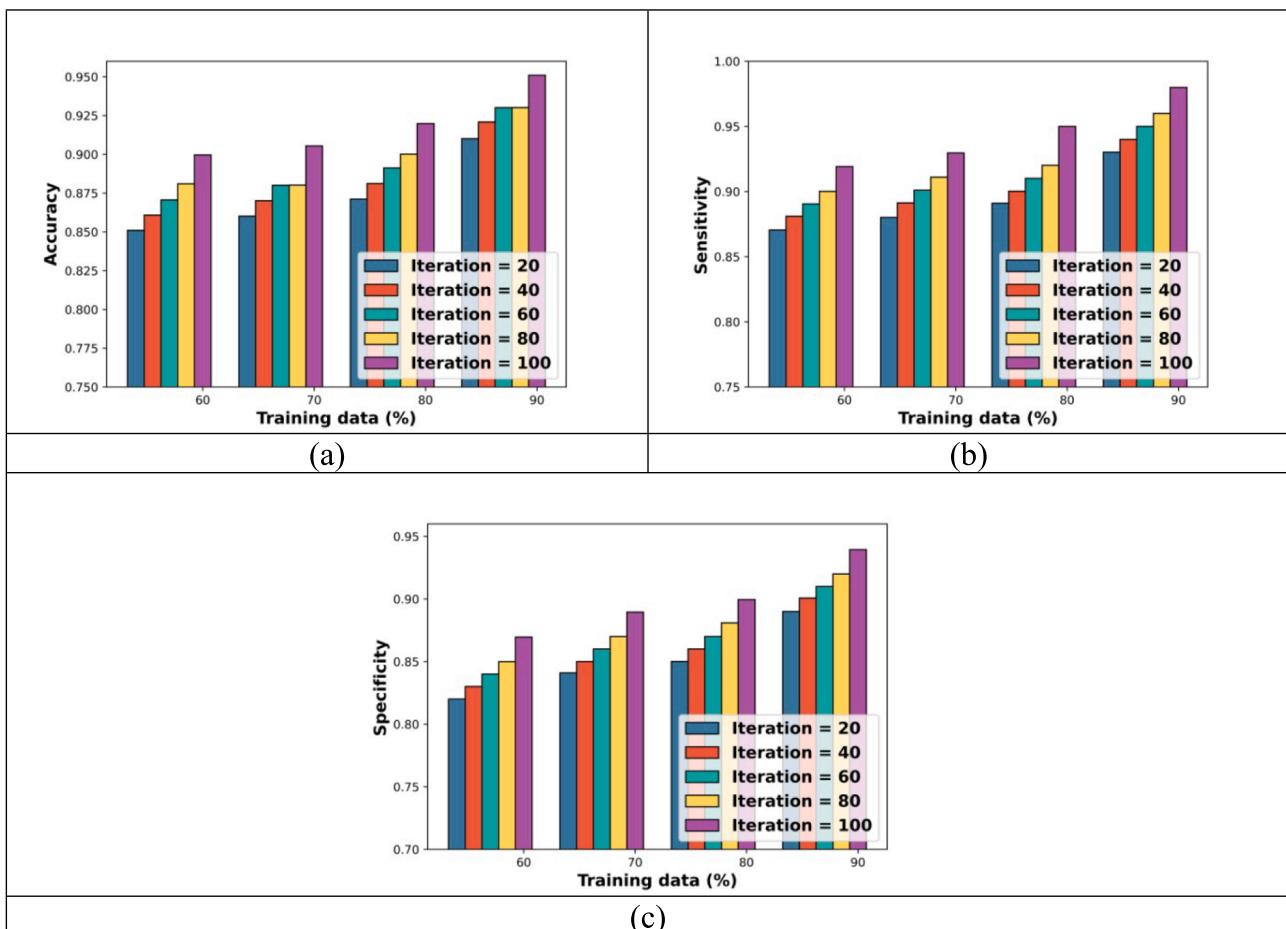
4.2.2. Sensitivity

Sensitivity metric is employed to compute true positive classification results and is mathematically defined as:

$$\text{Sensitivity} = \frac{A}{A + D} \quad (5)$$

4.2.3. Specificity

Specificity is a metric used to calculate true negative classification results, and its mathematical representation is:

**Fig. 6.** Performance analysis of proposed model by varying training percentage (a) Accuracy(b) Sensitivity (c) Specificity.

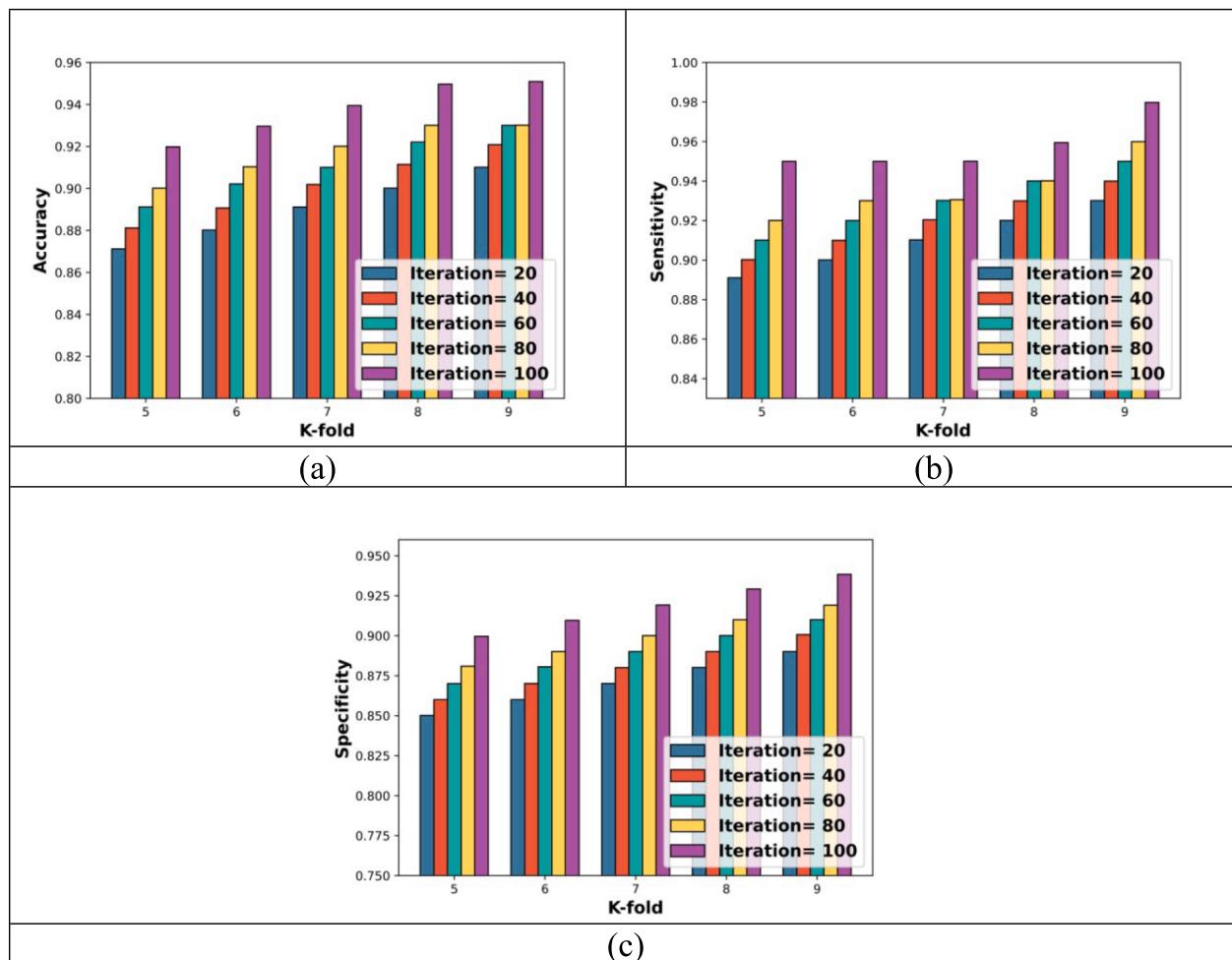


Fig. 7. Performance analysis of proposed model by varying k-fold (a) Accuracy (b) Sensitivity (c) Specificity.

$$\text{Specificity} = \frac{B}{B+C} \quad (6)$$

4.3. Performance analysis

Fig. 5. demonstrates the confusion matrix. Fig. 6 demonstrated performance analysis of Hybrid SqueezeNet-Inception V3 model while altering training percentage to assess changes in sensitivity, accuracy, and specificity. Fig. 6 (a) exhibited analysis of approach regarding accuracy. Here, accuracy of model was 0.920 for iteration 40 and it was maximized to 0.950 for iteration 100 when training percentage was 90. Fig. 6 (b) exhibited analysis of model regarding sensitivity. While training percentage was 90, sensitivity of model was 0.940 for iteration 40 and it was increased to 0.979 for iteration 100. Fig. 6 (c) shows analysis of model regarding specificity. The specificity of technique was 0.910 for iteration 60 and it was raised to 0.939 for iteration 100 while training percentage was 90.

Fig. 7 exhibited performance analysis of proposed Hybrid SqueezeNet-Inception V3 model based on sensitivity, accuracy, and specificity while altering the k-fold. Fig. 7 (a) shows analysis of model concerning accuracy. The accuracy of model was 0.900 for iteration 20 and it was increased to 0.949 for iteration 100 while k-fold was 8. The analysis of model regarding sensitivity is exhibited in Fig. 7 (b). When the k-fold was 7, sensitivity of model was 0.920 for iteration 40 and it was maximized to 0.950 for iteration 100. Fig. 7 (c) shows analysis of model regarding specificity. When K-fold was 7, specificity of the technique was 0.870 for iteration 20 and it was maximized to 0.919 for iteration 100.

4.4. Comparative analysis

Fig. 8 illustrates comparative analysis of models while altering training percentages for sensitivity, accuracy, and specificity. Fig. 8 (a) demonstrates an analysis of techniques regarding accuracy. Here, accuracy of the Hybrid SqueezeNet-Inception V3 was 0.919, whereas, the GAN, CNN and RetinaNet were 0.881, 0.891, and 0.900 when training percentage was 80. Fig. 8 (b) shows analysis of models in terms of sensitivity. Here, sensitivity of the Hybrid SqueezeNet-Inception V3 was 0.95, whereas, the GAN, CNN and RetinaNet were 0.890, 0.900, and 0.910 when training percentage was 80. The analysis of the models in terms of specificity is shown in Fig. 8 (c). Here, specificity of the Hybrid SqueezeNet-Inception V3 was 0.889, whereas, the GAN, CNN and RetinaNet were 0.840, 0.850, and 0.861 when training percentage was 70.

Fig. 9 illustrates an analysis of schemes while altering k-fold for sensitivity, accuracy, and specificity. Fig. 9 (a) shows analysis of techniques concerning accuracy. Here, the accuracy of the Hybrid SqueezeNet-Inception V3 was 0.949, whereas, the GAN, CNN and RetinaNet was 0.909, 0.915, and 0.925 when k-fold was 8. Fig. 9 (b) exhibited an analysis of models regarding sensitivity. When k-fold was 7, sensitivity of Hybrid SqueezeNet-Inception V3 was 0.95, whereas, the GAN, CNN and RetinaNet were 0.910, 0.921, and 0.930. The analysis of techniques regarding specificity was shown in Fig. 9 (c). When K-fold was 7, the specificity of the Hybrid SqueezeNet-Inception V3 was 0.919, whereas, the GAN, CNN and RetinaNet was 0.870, 0.880, and 0.890.

Fig. 10. demonstrates the analysis regarding intensity profile analysis. Here, pixel position is stated with horizontal axis and intensity is stated with vertical axis. The graph exhibits fluctuations in intensity

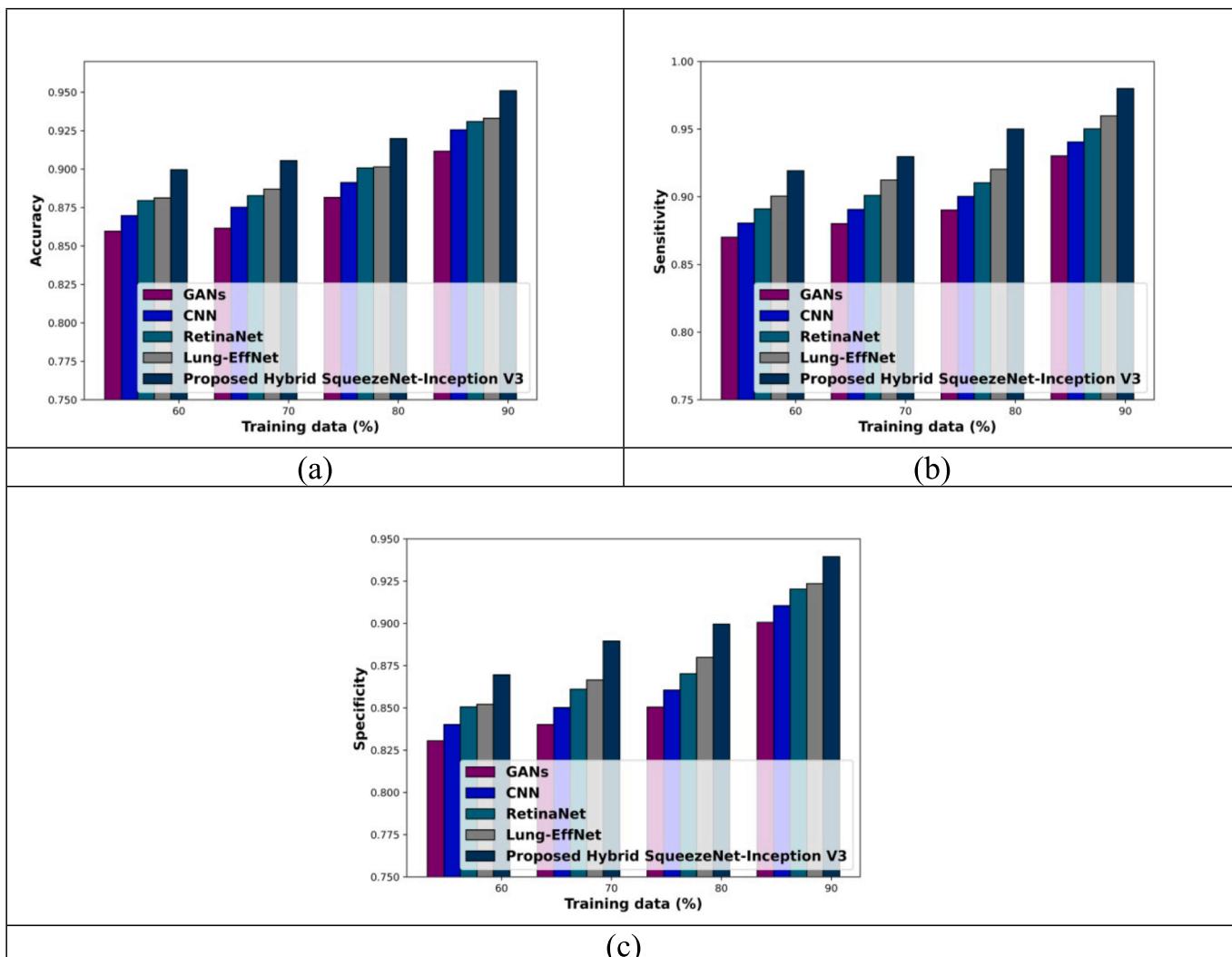


Fig. 8. Analysis of the models while altering training percentage (a) Accuracy (b) Sensitivity (c) Specificity.

with numerous prominent spikes indicating points of higher intensity.

4.5. Discussion

In Table 2, a comparison of models was presented, with variations in K-fold and training percentage taken into account. For training percentage, the proposed model was 4 % better than the GAN model for accuracy; proposed model was 3 % better than the CNN model for specificity; proposed model was 3 % better than the RetinaNet model for sensitivity. It was observed that proposed approach outperforms traditional methods regarding accuracy, specificity, and sensitivity when considering different K-fold values and training percentages. For K-fold, the proposed model was 4 % better than the GAN model for accuracy; proposed model was 3.1 % better than the CNN model for specificity; proposed model was 3.2 % better than the RetinaNet model for sensitivity. Table 3 demonstrates the computational analysis regarding training percentage and K-fold. Here, the proposed model attains less computational complexity of 686 for training percentage and 679 for K-fold.

5. Conclusion

This paper develops a novel Squeeze-Inception V3 with SMA-based CNN features to recognize lung cancer disease. The approach comprises phases involving pre-processing of Chest CT-Scan images, feature

extraction, and detection. In pre-processing stage, main aim was to improve the quality of raw Chest CT-Scan images using Color space conversion, data augmentation, resizing, and normalizing. The pre-processed outputs were fed to feature extraction, where CNN that was trained by the SMA was utilized for feature extraction. Subsequently, features were fed to the detection phase to detect lung diseases. The detection of normal/abnormal was performed by utilizing the proposed Squeeze-Inception V3, which was modeled by combining the SqueezeNet and Inception v3 architectures. At last, the experimental analysis stated that proposed model obtained maximal specificity, accuracy, sensitivity, and with the rate of 94 %, 95 %, and 98 %. In this study, only the detection of normal and abnormal labels is successfully performed. Nonetheless, future work could extend the classification task to include four classes: Adenocarcinoma, large cell carcinoma, normal, and squamous cell carcinoma. This extension could involve training a novel deep learning model by an optimization algorithm to handle the classification phase effectively across all four classes. This approach aims to enhance the model's ability to differentiate between different types, thus enhancing diagnostic accuracy and applicability in clinical settings.

CRediT authorship contribution statement

Geethu Lakshmi G: Funding acquisition. P. Nagaraj: .

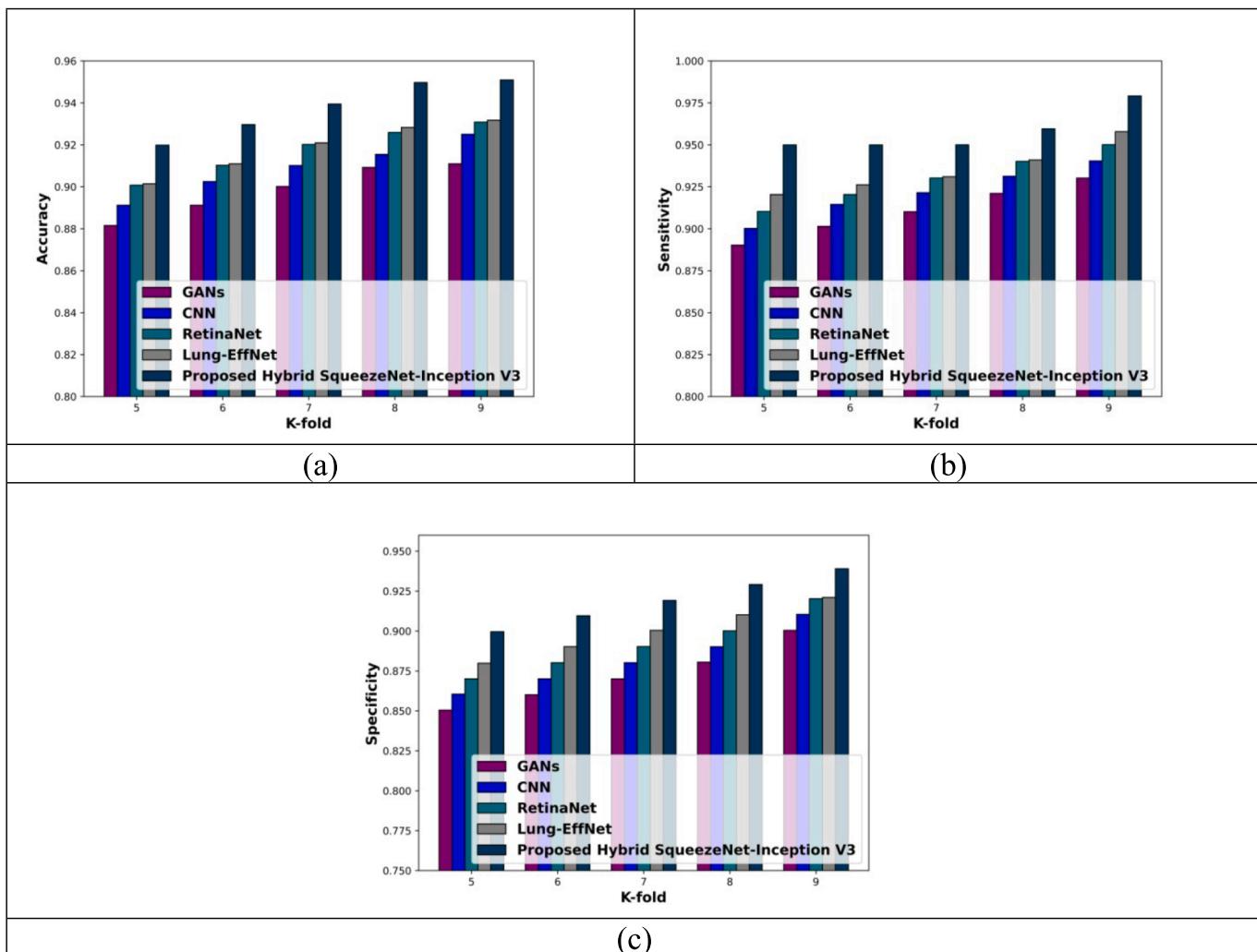


Fig. 9. Analysis of models while altering k-fold (a) Accuracy (b) Sensitivity (c) Specificity.

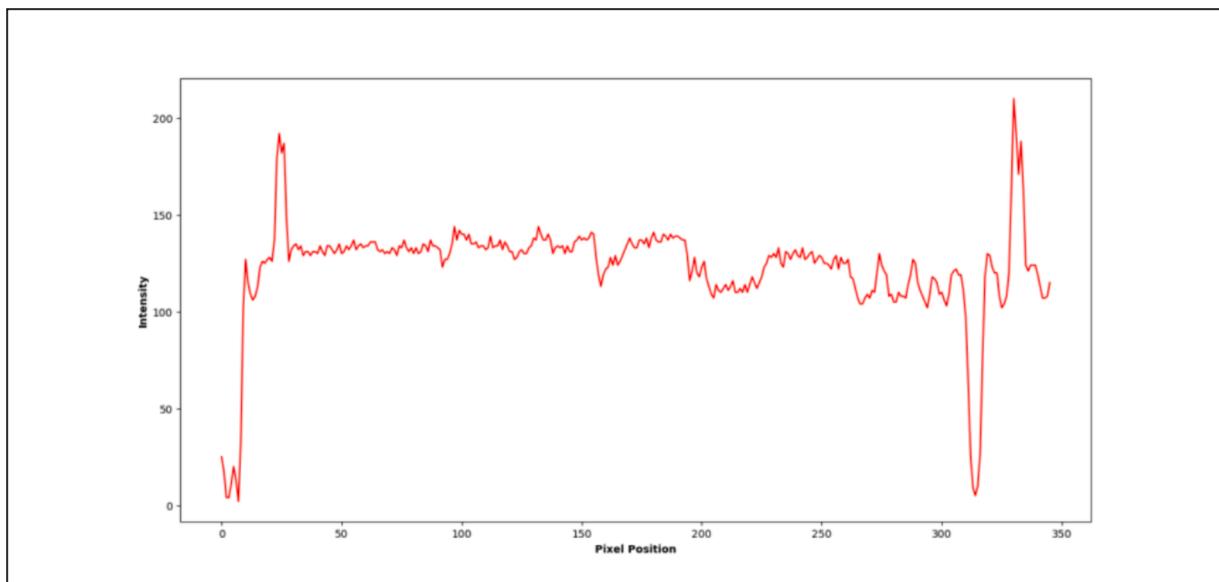


Fig. 10. Analysis regarding intensity profile analysis.

Table 2

Comparative Discussion regarding training percentage and K-fold.

Methods	Training percentage				
	GAN	CNN	RetinaNet	Lung-EffNet	Proposed model
Accuracy	0.912	0.926	0.931	0.933	0.951
Specificity	0.900	0.910	0.920	0.960	0.939
Sensitivity	0.930	0.940	0.950	0.92349	0.980
K-fold					
Accuracy	0.911	0.926	0.931	0.932	0.951
Specificity	0.901	0.910	0.920	0.958	0.939
Sensitivity	0.930	0.940	0.950	0.921	0.980

Table 3

Computational analysis regarding training percentage and K-fold.

Methods	Training percentage				
	GAN	CNN	RetinaNet	Lung-EffNet	Proposed model
970	899	828	757	686	
K-fold					
1017	965	853	731	679	

Declaration of competing interest

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.

Data availability

The authors do not have permission to share data.

References

- [1] T.J. Doyle, A.S. Patel, H. Hatabu, M. Nishino, G. Wu, J.C. Osorio, M.F. Golzarri, A. Traslosheros, S.G. Chu, M.L. Frits, C.K. Iannaccone, Detection of rheumatoid arthritis-interstitial lung disease is enhanced by serum biomarkers, *Am j Respir Crit Care Med.* 191 (12) (2015) 1403–1412.
- [2] R. Mahum, A.M. AlSalman, Lung-retinanet: Lung cancer detection using a RetinaNet with multi-scale feature fusion and context module, *IEEE Access* (2023).
- [3] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, CNN features off-the-shelf: an astounding baseline for recognition, In Proc IEEE conf comput vision pattern recognit workshops. (2014) 806-813.
- [4] S. Lum, P. Gustafson, H. Ljungberg, G. Hülskamp, A. Bush, S.B. Carr, R. Castle, A. F. Hoo, J. Price, S. Ranganathan, J. Stroobant, Early detection of cystic fibrosis lung disease: multiple-breath washout versus raised volume tests, *Thorax.* 62 (4) (2007) 341–347.
- [5] P. Yadav, N. Menon, V. Ravi, S. Vishvanathan, Lung-GANs: Unsupervised representation learning for lung disease classification using chest CT and X-Ray images, *IEEE Trans Eng Manage.* 70 (8) (Aug 2023) 2774–2786.
- [6] A. Asuntha, A. Srinivasan, Deep learning for lung Cancer detection and classification, *Multimedia Tools and Appl.* 79 (2020) 7731–7762.
- [7] S.L.F. Walsh, L. Calandriello, M. Silva, N. Sverzellati, Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study, *The Lancet Respir Med.* 6 (11) (2018) 837–845.
- [8] K. Sripornt, C.-F. Tsai, C.-E. Tsai, and P. Wang, Analyzing lung disease using highly effective deep learning techniques, In *Healthcare.* 8(2) (2020) 107. MDPI.
- [9] P. Podder, S.R. Das, M.R.H. Mondal, S. Bharati, A. Maliha, M.J. Hasan, F. Piltan, Lddnet: a deep learning framework for the diagnosis of infectious lung diseases, *Sens.* 23 (1) (2023) 480.
- [10] P. Yadav, N. Menon, V. Ravi, S. Vishvanathan, Lung-GANs: unsupervised representation learning for lung disease classification using chest CT and X-ray images, *IEEE Trans Eng Manage.* (2021).
- [11] B. Liu, H. Yu, X. Zeng, D. Zhang, J. Gong, L. Tian, J. Qian, L. Zhao, S. Zhang, R. Liu, Lung cancer detection via breath by electronic nose enhanced with a sparse group feature selection approach, *Sens Actuators B: Chemical.* 339 (129896) (2021).
- [12] M. Togacar, B. Ergen, Z. Comert, Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks, *Biocybern Biomed Eng.* 40 (1) (2020) 23–39.
- [13] S. Ram, W. Tang, A.J. Bell, R. Pal, C. Spencer, A. Buschhaus, C.R. Hatt, M. P. diMaglione, A. Rehemtulla, J.J. Rodriguez, S. Galban, Lung cancer lesion detection in histopathology images using graph-based sparse PCA network, *Neoplasia.* 42 (2023) 100911.
- [14] O. Ozdemir, R.L. Russell, A.A. Berlin, A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans, *IEEE Trans Med Imaging* 39 (5) (2019) 1419–1429.
- [15] L. Pham, H. Phan, R. Palaniappan, A. Mertins, I. McLoughlin, CNN-MoE based framework for classification of respiratory anomalies and lung disease detection, *IEEE J Biomed Health Inf.* 25 (8) (2021) 2938–2947.
- [16] R.R. Shivwanshi, N. Nirala, Hyperparameter optimization and development of an advanced CNN-based technique for lung nodule assessment, *Phys Med Biol.* 68 (17) (2023) 175038.
- [17] R.R. Shivwanshi, N. Nirala, Quantum-enhanced hybrid feature engineering in thoracic CT image analysis for state-of-the-art nodule classification: an advanced lung cancer assessment, *Biomed Phys & Eng Express.* 10 (4) (2024) 045005.
- [18] S. Li, H. Chen, M. Wang, A.A. Heidari, S. Mirjalili, Slime mould algorithm: A new method for stochastic optimization, *Future Gener Comput Syst.* 111 (2020) 300–323.
- [19] Z. Bi, and P. Cao, Color space conversion algorithm and comparison study, In *Journal of Physics: Conference Series.* 1976(1) (2021) 012008. IOP Publishing.
- [20] M. Alnowami, E. Taha, S. Alsebaei, S.M. Anwar, A. Alhawsawi, MR image normalization dilemma and the accuracy of brain tumor classification model, *J Radiat Res Appl Sci.* 15 (3) (2022) 33–39.
- [21] K. Dharavath, G. Amarnath, F. A. Talukdar, and R. H. Laskar, Impact of image preprocessing on face recognition: A comp anal, In *2014 Int Conf Commun Signal Proc.* (2014) 631–635. IEEE.
- [22] D. Kumar, and V. Kukreja, N-CNN based transfer learning method for classification of powdery mildew wheat disease, In *2021 int conf Emerging Smart Comput Inf (ESCI).* 707–710 (2021), IEEE.
- [23] M. Ragab, I. Katib, S.A. Sharaf, F.Y. Assiri, D. Hamed, A.L. Abdullah, Self-upgraded cat mouse optimizer with machine learning driven lung cancer classification on computed tomography imaging, *IEEE Access* (2023).
- [24] X. Zhu, Z. Feng, Y. Fan, J. Ma, A multiple-blockage identification scheme for buried pipeline via acoustic signature model and SqueezeNet, *Meas.* 202 (2022) 111671.
- [25] F. Ucar, D. Korkmaz, COVIDagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images, *Med Hypotheses.* 140 (2020).
- [26] G. Meena, K.K. Mohbey, S. Kumar, Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach, *Int J Inf Manage Data Insights.* 3 (1) (2023).
- [27] Chest CT-Scan images Dataset, https://www.kaggle.com/datasets/mohamedha_nyy/chest-ctscan-images, accessed on July 2022.
- [28] R. Raza, F. Zulfiqar, M.O. Khan, M. Arif, A. Alvi, M.A. Iftikhar, T. Alam, Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images, *Eng Appl Artif Intell.* 126 (2023) 106902.