International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

Lung Disease Classification using Dense Alex Net Framework with Contrast Normalisation and Five-Fold Geometric Transformation

Poonam Rana[1], Vineet Sharma[2], Pradeep Kumar Gupta[3]
[1]Scholar Dr.A.P.J. Abdul Kalam Technical University,
poonam.rana1988phd@gmail.com
[2]Krishna Institute of Engineering and Technology, MuradNagar, U.P.,201206,
Vineet.sharma@kiet.edu
[3]Jaypee university of Information Technology, Waknaghat, (H.P.),
pkgupta@ieee.org

Abstract— lung disease is one of the leading causes of death worldwide. Most cases of lung diseases are found when the disease is in an advanced stage. Therefore, the development of systems and methods that begin to diagnose quickly and prematurely plays a vital role in today's world. Currently, in detecting differences in lung cancer, an accurate diagnosis of cancer types is needed. However, improving the accuracy and reducing training time of the diagnosis remains a challenge. In this study, we have developed an automated classification scheme for lung cancer presented in histopathological images using a dense Alex Net framework. The proposed methodology carries out several phases includes pre-processing, contrast normalization, data augmentation and classification. Initially, the pre-processing step is accompanied to diminish the noisy contents present in the image. Contrast normalization has been explored to maintain the same illumination factor among histopathological lung images next to pre-processing. Afterwards, data augmentation phase has been carried out to enhance the dataset further to avoid over-fitting problems. Finally, the Dense Alex Net is utilized for classification that comprises five convolutional layers, one multi-scale convolution layer, and three fully connected layers. In evaluation experiments, the proposed approach was trained using our original database to provide rich and meaningful features. The accuracy attained by the proposed methodology is93%, which is maximum compared with the existing algorithms

Index terms: Dense AlexNet, Deep learning, geometric transformation, contrast normalization, median filter, Gaussian filter, adenocarcinomas, squamous cell carcinomas.

I. Introduction

In India, one of the leading causes of death is lung cancer, which is rapidly increasing with a dramatic upsurge in cigarette smoking. Lung cancer is a malignant tumour characterized by uncontrolled cells' growth in the lung tissue [1]. Primary lung cancers are divided into two main types, adenocarcinoma and squamous cell carcinoma. Adenocarcinoma and squamous cell carcinoma are challenging to differentiate precisely in their morphological features, require immune-histochemical evaluation [2, 3]. Recently, digitized histopathological images have been rapidly growing due to its advanced features. The treatment of this disease varies from patient to patient based on nodule growth and lung changes. Therefore, providing treatment in advance helps to avoid spreading of disease from one part of the body to others. Thus the early diagnosis of the type of lung cancer is the primary goal for researchers and medical Practitioners. Classification of histopathological images based on tumour subtypes is the primary pathological imaging for classifying lung cancer diseases [4-6]. Recently, enhancements have been made in the examination process with the help of digital imaging devices. Additionally, AI methods are utilized in the biomedical field [7, 30].

Histopathological images consist of many tissues and cells. Thousands of tissue and cells areas must be analyzed to interpret these tumour cells' interpretation in whole slide images. In some prediction, misdiagnosis was caused by negative factors like less experience of pathologist and fatigue [8, 9]. To avoid these negative factors, Deep Learning methods have been used for the decision making process. Advantages of Deep Learning algorithms have collaborated with these image classification methods increased the accuracy in prediction and classification of Lung Cancer disease [10]. This is the reason for a misleading background texture and solved it by using simple image processing methods. Image pre-processing techniques on various Datasets have been employed to remove unwanted noise from the histopathological Images .Hand-crafted feature extraction methods are useful in pre-processing and had been used previously [11]. Automatic feature extraction algorithms

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

have risen increasingly focused by researchers for an appropriate and time-consuming diagnosis. As they extract features, every data in the image is necessary.. In this case, there is a problem, automatic feature extraction automatically generate features, and hence each data in the image is essential [12]. A complex issue in managing such disease is that number of incorrectly diagnosed patients and if misdiagnosed leads to mistreatment. Although all the works mentioned above help analyse and classify lung diseases to an extent, there is always a trade-off between accuracy and computational complexity. To overcome this issue, increase the accuracy, reduce the computational cost, feature selection approach should be employed in the Computer Aided Diagnosis System [13, 14].

Numerous pathologists have focused on Computer-aided diagnosis to investigate on lung cancer and its types. Also detecting multiple objects and classifying tumour from massive datasets is difficult [21-23]. Based on the survey conducted by big data research, deep learning techniques and Architectures can enhance the diagnostic accuracy rate by identifying and making decisions in the biomedical field Deep networks learn in-depth features and identify a necessary object by utilising a series of deconvolution and convolution layers. With the help of deep learning Convolutional Neural Network Algorithm, Histopathological characteristics of each tumour tissue affected by lung cancer can be separated into lung adenocarcinoma (ADC) and Squamous cell carcinoma (SCC) [17]. These are two major types of histopathological images, but their biological characteristics differ. Even though prognostic markers are available for adenocarcinoma, this study utilises deep learning approaches to classify lung cancer images and further undergo treatment for severe lung cancers [17].

The contribution of the proposed research is illustrated as follows

- To increase the performance of lung cancer Prediction and Classification accuracy, pre-processing is done for the considered lung cancer histopathological image dataset. With the pre-processing step's, image information gets enhanced by smoothening the image and reduce the noise from all the images in the Histopathological Images Datasets.
- To retain uniform intensity, contrast normalization is carried out in the proposed methodology to get the normalised image This in turn, helps to improve the learning rate of the network model.
- To extend training images for avoiding over-fitting issues, data augmentation phase is done. This helps to get a deeper relation among pixels in histopathological lung images.
- To obtain features of different scales, multi-scale convolution layer is applied in the proposed Dense Alex net Framework. It integrates PRelu activation function for avoiding information loss and strengthen the overall network model.

The organisation of the manuscript are as follows, Section 2 discuss related works concerning proposed model. Section 4 briefs about the proposed model and Section 5 details the simulation results. Finally, section 6 concludes the whole paper.

II. Literature Review

In this section, some of the recent related works on lung cancer analysis are presented, the procedure involved in the previous study, and its limitations.

Alzubi et al [15] introduced an approach for lung cancer disease and was named Weight Optimized Neural Network with Maximum Likelihood Boosting (WONN-MLB). This technique utilised big data for analysis and was carried out by two stages: feature selection and classification. The relevant lung disease attributes are selected through combined pre-processing methods such as Newton Rsphsons Maximum Likelihood and Minimum Redundancy (MLMR) for minimising the time of classification performance. Then the presented technique ensemble classification based boosted weighted optimised neural network algorithm was proposed. The developed classification method helps to enhance diagnosis accuracy by reducing false occurrence. But the feature selection approach leads computation complexity to the system. In 2014, Bhuvaneswari et al. [16] have exhibited a feature extraction approach for CT lung images. With the integration of both Walsh Hadamard transform features and Gabor filter, the authors invented a fusion-based method using Median Absolute Deviation (MAD) model. The authors proposed the system by carrying three stages, namely pre-processing, feature extraction and classification. Initially, the pre-processing steps applied to the Histopathological Images Datasets where all the noise should be removed from the images and then feature extraction approach involves the presented fusion technique. Then genetic algorithm was utilised to choose the relevant set of features from the dataset.Classification algorithms mainly K nearest neighbour (KNN), decision tree and Multi-layer perceptron Neural Networks (MLP-NN) were utilised to perform the classification of lung disease. Karthiga*et al* [18] had presented an evolutionary algorithm termed Accelerated Wrapper based Binary Artificial Bee Colony (AWB-ABC) for an efficient feature selection. It enhanced the performance of Naive bayes classifier for different types of Cancer. With LUNA16 as input dataset, after pre-processing this data to remove noise and improve image, it was done. From this morphological features were extracted and that cause cancer were differentiated.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

However, these classifications were done based on naive byes method as it has comparatively low performance. Suresh *et al.* [19] had presented a technique for automated tumour stage classification of pulmonary lung nodules. It utilised an end-to-end Deep Convolutional Neural Network (DCNN). From an open-source lung cancer consortium, the dataset was obtained and segmented into 52 x 52-pixel nodule depending on radiologist's opinion and Ground truth values. It analysed and extracted self-learned salient features from these segmented images and fed it to a trained DCNN for classification. However, tumour patterns were alone used for classification of images as non-cancerous and cancerous hence appropriate classification was unable to achieve. [20] had presented a method for classification and detection of diseased tissue and healthy lungs with Computed Tomography (CT). This method used a deep learning approach for the identification of lung tissue texture signatures. An interstitial Lung Disease (ILD) dataset was used in this analysis, based on the fusion of Rizez, pre-processing was done. Computing features based on the Convolutional Neural Network (CNN) and fine-tuning are both non-parametric and deep learning approaches. This method analysed the combination of joint softback model classification for comparison of previous and late features. Li *et al.* [24] had presented an adaptive multinomial regression with a sparse overlapping group to perform classification of grouped gene selection for lung cancer.

Lakshmanaprabu*et al* [25] had analysed lung images obtained from CT scan in assistance with Optimal Deep Neural Network (ODNN) and Linear Discriminative Analysis (LDA). This analysis extracted deep features from this lung images, and their dimensionality was reduced with the help of LDR for classification of lung tissues. By applying ODNN to lung images, optimisation was done using Modified Gravitational Search Algorithm (MGSA) to predict lung cancer. However, this analysis considered lung images obtained from CT scan alone. Liu *et al.* [26] had presented a method for lung cancer recognition based on image quality assessment. It classified sick lung images from healthy ones. Data from Low-dose CT scan was obtained with the whole iterative combination method.

Again incorporating both low and high-level features for the extraction of in-depth features from a dataset. This method used a CNN based feature classification. However, this method required noise level extraction and also artefacts was increasing in amount. Liu *et al.* [28] had introduced a FIG concept under the basis of genetic optimisation and fisher standards for categorising a feature selection model. It determined fisher standards by evaluating different subset feature. By incorporation of genetic algorithm to fisher standards for each characteristic enhancement for selecting optimal subset features. By utilising FIG model, this predicted different lung disease images from obtained CT image. However, this technique carries bag-of-visual word based on a histogram of oriented gradients by wavelet transform based features. Also, this method used different techniques that increased diagnosis time. Hawkins *et al.* [29] had presented an analysis of applying 3-D elements to the Lung cancer CT images. It offered a prognostic information classification using Analog to Digital Converter (ADC). A large set of data non-small cell lung cancer was categorised into lung tumour and its subtype. It helped for the enhancement of the survival rate of patients. However, this research only focussed on ADC type lung disease obtained from CT type screening images. Petousis*et al* [30] have investigated lung cancer screening using LDCT based lung cancer dataset to achieve an effective false-positive rate. It was done by training dynamic Bayesian Network as a model, and it utilised inverse reinforcement learning for finding lung cancer tissue based on expert decisions. It also combined multiple machine learning methods with Markov decision-making to improve simultaneous lung cancer detection for diagnosis purpose. However, final decision was made based on expert opinion, and hence it requires lengthy diagnosis time.

Performance of these existing systems depends heavily on the previous step's accuracy, which is a primary drawback. Abnormalities of lungs are assessed through the computer-aided procedure and diagnostic reports sent to a specialist to support them for planning treatment accordingly. Traditionally hand-crafted feature extraction was done and based on expert opinion. These methods made decisions based on this opinion over patients. Since the specialists and experts in some cases were unable to predict infected region until the patient reach a severe stage. Previously existing diagnosis methods for distinguishing between lung cancer images failed to identify little knob that would lead to the cancerous lung in the future. Due to low accuracy and ineffective decision making, existing techniques were unable to provide enhanced classification. These limitations motivated a lot to promote an effective algorithm for developing automatic histopathological lung image classification system based on deep learning.

III. Research Methodology

The growing demand for lung disease classification using histopathological images leads to the development of new innovation for enhancing the lung disease classification accuracy Aiming for this objective, deep learning technologies and algorithms proposed to offer an automatic detection system of lung nodules to save human beings lives with early detection and diagnosis mechanisms. Therefore in this research, a novel Dense AlexNet framework is designed to classify dissimilar lung diseases.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

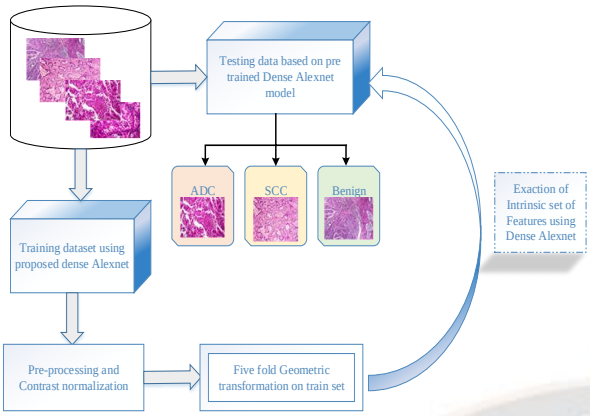


Figure 1: Overall proposed Architecture

The deep learning classification model can process raw sets of input data and automatically extract the deep features without the hand-crafted feature extraction approach. It can perform without implementing the pre-processing tasks, but the image must be normalised with suitable pre-processing procedures to augment the learning procedure quickly. For this purpose, the input lung histopathological images are initially fed into pre-processing steps. To maintain the lung histopathological images with similar illumination factors, the contrast normalisation phase is carried out, which augments the adaptability of the proposed classification technique. While using the deep learning approach it is necessary to enrich the training dataset for showing the variation of images. Hence five-fold geometric transformation is performed on training set images to improve the classification accuracy for proposed Dense AlexNet framework. Finally, designed Dense AlexNet framework tends to classify three lung cancer stages: ADC, SCC, and benign.

3.1. Image enhancement using pre-processing procedures

Histopathological image analysis of the lungs is challenging while leading directly to machine learning techniques and the pathologist to diagnose it. In the case of a diagnosis performed by pathologists, it takes about an hour for a single image to analyse the whole-slide tissue. When it comes to machine learning methods, the desired performance is not achieved by many classification methods. The result is not up to the mark due to manual crated feature extraction and unsuitable learning rates. The causes of these problems are of image possessing high disruption and resolution factors. Hence subjecting these raw images directly to the learning approach means that the performance decreases in terms of the training process and learning capability. Thus, to resolve these challenges, efficient pre-processing steps are incorporated in the proposed methodology to accelerate the network learning rate and eliminate disturbing factors. Hence for removing the fluctuations from input lung histopathological images, pre-processing procedures are applied to get the pre-processed image.

Typically the input image size followed by conventional AlexNet framework is in the form of 256×256 pixel size. The proposed Dense AlexNet architecture follows a similar pixel size. Hence before feeding the input lung histopathological lung images to learning strategies, image resizing is carried out and set to pixel size of 256×256. Then the resized lung histopathological image $J(y,z)$is applied with Gaussian Filter for smoothening the image by diminishing the noise present in it. It also helps to sharpen the object contained in the image. The following notation mathematically expresses it.

$$I_1(y,z) = J(y,z) * \frac{e^{\frac{-(y^2+z^2)}{2\sigma^2}}}{\sum_y \sum_z e^{\frac{-(y^2+z^2)}{2\sigma^2}}}, \quad \sigma = 2 \qquad (1)$$

Here the sigma value of Gaussian filter is chosen to be as 2. The reason behind selecting sigma value as 2 is that the smaller sigma values are limited to only the local regions alone. In contrast, suppose the chosen sigma value is considerable, then the intracellular grey values affect the background region's overall change. Then for a background image, the median value of pixels is estimated. After computing the pixels' median value, it is subtracted mathematically from every pixel in the obtained smoothened image. The following equation mathematically expresses it.

$$I_2(y,z) = I_1(y,z) - median\big(I_1(y,z)\big) \qquad (2)$$

This procedure helps to eradicate the unnecessary background noises present in the image. Then the proposed pre-processing module tends to find the region possessing a higher impact of noise. This image is removed from the original image to compensate for unexpected colour glare and noise enclosed in the image. The following notation can mathematically express it.

$$I_3(y,z) = J(y,z) - I_2(y,z) \qquad (3)$$

The obtained image is then sharpened with a sharpening filter that leads to highlights the fine details and edges of an image. The following notation can mathematically express it.

$$I_4(y,z) = I_3(y,z) - f_{smooth}\big(I_3(y,z)\big) \qquad (4)$$

In the last step of the pre-processing procedure, the noise present in the $I_4(y,z)$sharpened image is removed using a two-dimensional median filter with 5×5 adjacency. It is expressed in the underneath equation.

$$I_5(y,z) = f_{median}[5 \times 5]\big(I_4(y,z)\big) \qquad (5)$$

The proposed image enhancement procedure is done via the proposed pre-processing technique by integrating various tasks.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

The followed pre-processed resultant image enhances the classification model's performance by diminishing the disruptive factor and effect of noisy contents present in the image. The pre-processing step is mandatory because the deep learning technique doesn't try to get meaningless features like the background changes etc. Due to these benefits, the learning rate of the proposed classification model is accelerated.

3.2. Contrast Normalisation for the pre-processed outcome

Generally, input lung histopathological images are captured under different scanners and may be prepared with other chemicals so that its illumination varies. This leads to microscopic images of different intensity and contrast that impede the simplification of performance and learning ability. Contrast Normalisation (CN) is applied to avoid the contrast variation among pre-processed images to overcome this challenge. The meanwhile-processed outcome is in the form of RGB histopathological image, and it is subjected to CN phase in a channel-wise manner. For every data, primarily mean and standard deviation are computed for three channels in CN phase. After calculating both, the mean value is subtracted from three channels in the attained pre-processed image and dividing it by estimated standard deviation. The following mathematical expression evaluates the CN phase.

$$l(y,z) = \frac{I_5(y,z)-\mu}{\sigma} \qquad (6)$$

Where $l(y,z)$ indicates the normalised image, $I_5(y,z)$ represents the pre-processed outcome, μ symbolizes the estimated mean value calculated for each channel and σdenotes the standard deviation for the RGB channel. Equation 6 is applied across three channels to get the normalised image. With the proposed contrast normalisation phase's help, the proposed model's adaptability is strengthened dependent on the target data values. Also, CN helps to diminish contrast variation during the classification model's training between images.

3.3. Five-Fold Geometric transformation

Five-fold data augmentation is carried out for the normalised lung histopathological images during the training of the models. By this procedure, over-fitting problems are avoided and enhance the model generalisation ability. For the smaller dataset, the performance of deep learning mechanism is hindered by over-fitting issue. Hence deep learning techniques perform better while extending the dataset using geometric transformation method. This transformation technique eases the pathologist's analysis process by augmenting the histopathological lung images with various sizes, angles, and orientations. Thus a large set of training images are obtained with less effort. Here five geometric transformation is applied for a single image selected. Thus each sample image is transformed to generate five additional samples in the training set. The proposed three fold geometric alteration states that altering the geometry of the normalised image by relocating the object to a new orientation and position. This helps to preserve the overall shape of an image. The five-fold geometric transformation encompasses scaling, horizontal flipping, vertical flipping, cropping, rotation which transforms the image through mapping around its centre of each image pixels $(y,z)$to $(\grave{y},\grave{z})$. The following mathematical transformation can do it.

$$\begin{pmatrix}\grave{y}\\ \grave{z}\end{pmatrix} = \begin{bmatrix} p\cos\theta & -\sin\theta \\ \sin\theta & q\cos\theta \end{bmatrix}\begin{pmatrix} y \\ z\end{pmatrix} \qquad (7)$$

Here p and $q$are employed to control the scaling performance and can be set randomly. Likewise, θ indicatesrotation angle which is set to 180 degrees in horizontal flipping and vertical flipping. It can also be selected randomly. Moreover, the final geometric alteration performed is cropping, whereas p and q parameters be viewed as piecewise operation. The following notation can represent it.

$$p = \begin{cases} 1, & y \le \grave{w} \\ 0, & y \ge \grave{w} \end{cases} \qquad (8)$$

$$q = \begin{cases} 1, & z \le \grave{h} \\ 0, & z \ge \grave{h} \end{cases} \qquad (9)$$

Equation (8) and (9)possesses the pixel in the range of$(\grave{w},\grave{h})$. Here $\grave{w}$denotes the width of the image after transformation. Similarly $\grave{h}$ symbolises the height of the image after transformation. With the mentioned transformation steps, the deep learning performance gets improved by attaining the maximum accuracy of lung disease recognition rate.

3.4. Lung disease classification using proposed Dense Alex Net Framework

Due to the severity of lung cancer, disease classification is performed using a novel deep learning strategy. The reason for choosing deep learning approaches is because of the complex structure of histopathological images. The classification model's input is normalised histopathological augmented images subjected to the proposed Dense AlexNet framework. The proposed deep learning approach can extract the advanced features automatically with high disease recognition rate. To improve the ability of feature extraction process, the proposed method employs a deep convolution layer primarily to the conventional AlexNet architecture. The network structure of traditional deep learning is very complex, and so it requires a large scale of lung histopathological images. Also, it is difficult to train a larger number of lung disease images in traditional classification models and are trained too deeply only when the amount of image count is small.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

To overcome this challenge, a novel Dense Alex Net classification technique is designed based on the backbone of traditional AlexNet model. This in terms enhances the performance of training time and learning capability of the network. The proposed Dense AlexNet architecture involves 5 numbers of convolution layers, 1 multi-scale convolution part and 3 numbers of fully connected layers. These layers are incorporated together for the progression of network learning rate and avoid network complexity, which helps to enhance the network's classification accuracy by diminishing over-fitting and under-fitting issues. The convolution layers employed here assures the extensive range of feature extraction. Similarly, multi-scale convolution layer accomplishes kernels for attaining features with varying scale images. The final convolution layer encompasses the previous layers' features and the last three fully connected layers perform the classification task. Each convolution layer is followed by the Batch Normalisation layer that enhances the simplification of the proposed framework. The multi-scale convolution layer is set before the last convolution layer in the network model. The layout of the proposed Dense AlexNet framework is displayed below:

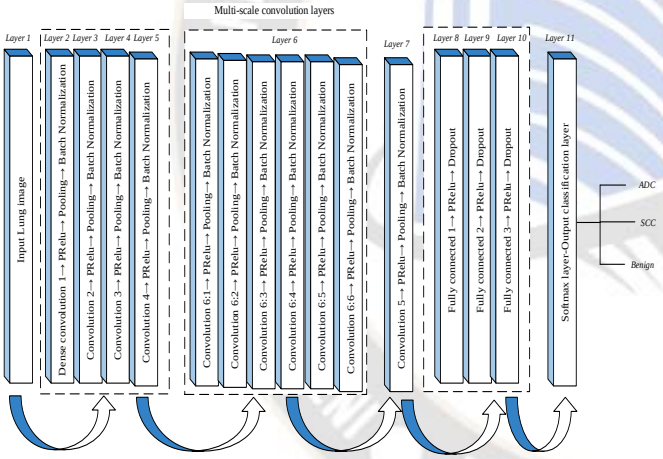


Figure 2: Layout of the proposed Dense AlexNet architecture

The detailed description of each layer is elaborated in the following sub-section.

3.4.1.Dense Convolution layer

The features like shape, size and position of lung disease are significantly varying in different phases of its growth, so they are changeable. Therefore it is necessary to extract a wider range of features for attaining the corresponding disease class. With this objective, some dense set of convolution layers are adopted in the first layer to enhance the network model's receptive field. The receptive field helps improve the network model's extraction ability and significant changes happening in the last layers of the network are also avoided is the added advantage. This, in turn, helps to compensate the loss function arisen in the classification model by performing a data sampling approach on feature maps. The receptive field is defined as the size of the region mapped on the original lung histopathological image through the pixel on the outcome of the feature map of each layer in the network model. The following mathematical notation expresses the receptive field of the network model.

$$s_{j+1}^2 = \left[(s_j - 1) + (2m + 1)\right]^2 \quad (10)$$

Herethe termss_j denotes the receptive field edge length of j^{th} layer, whereas $m$indicates the coefficient of the extended dense convolution layer.

3.4.2. Multi-scale Convolution layer

It is necessary to achieve maximum accuracy in the classification of different types of lung disease. Typically deep learning approaches improves the classification accuracy by augmenting the parameters and depth of the network. Still, only increasing the network depth faces computational complexity and over-fitting issues. To overcome from this difficulty, a multi scale convolution layer is added to the network model. The convolution kernels are introduced in this model. If only the convolution layer gets enlarged with different sizes, the layer integrates multiple parameters and makes the network complex and inefficient. Thus convolution kernels are utilised in this multi-scale convolution layer. Here the same feature map uses multiple convolution kernels that possess varying sizes to obtain the advanced features with different scales. Hence it is better to utilise multiple convolution kernels with varying sizes than single convolution layer. Max-pooling operation is used here to capture a different set of features. Max pooling is performed to choose the pooling region's maximum value by reducing the number of parameters when the images are too large. Hence this function is used to implement dimension reduction. Even though the dimensionality of feature maps is reduced, it retains the important features for further processing.

3.4.3. Batch Normalisation

The lung disease features are constant and complex; hence, the network model's learning strategy seems to be complicated. As the network structure gets deeper in this model, the distribution of features accomplishing in the hidden layer has suffered from fluctuations and significant changes. This leads to a negative impact on network stability. In this manuscript, the batch normalisation approach is utilised to normalise each layer's feature information by carrying a standard deviation of 1 and a mean of 0. These determined values help to train the proposed deep learning mechanism more stably and straightforwardly. It helps to improve the capability of network generalisation. The batch normalisation procedure evaluates each batch sample mean and standard deviation, which is mathematically solved by the following equations:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

$$\mu = \frac{1}{m}\sum_{j=1}^{m} y_j \quad (11)$$

$$\sigma = \frac{1}{m}\sum_{j=1}^{m}(y_j - \mu)^2 \quad (12)$$

The term μ indicates the mean of each batch samples and σ denotes the standard deviation of the batch sample y. Based on the calculation of both mean and standard deviation, the batch normalisation algorithm is mathematically expressed as follows:

$$\hat{y}_j = \frac{y_j - \mu}{\sqrt{\sigma^2 + \varepsilon}} \quad (13)$$

Here ε symbolises constant term which helps to avoid when the standard deviation term leads to zero so that error gets diminished.

3.4.4. Activation Function

The intensity range of histopathological image pixels may vary due to interference caused by illumination changes. In this situation, the generally sigmoid function is performed in the network model to suppress gradient change in the image region. But this sigmoid function possesses low convergence measure. Compared to sigmoid and tanh operation, Relu function achieves faster convergence rate. This, in turn, causes the gradient of the network to zero, which stops the system performance. Hence to overcome this challenge, a novel activation function named PRelu (Parametric Rectified Linear Unit) is utilised in the proposed Dense AlexNet model. The following notation can mathematically express it.

$$PRelu(y) = \begin{cases} y & if\ y > 0 \\ ay & if\ y \leq 0 \end{cases} \quad (14)$$

Where y denotes the obtained outcome of the neuron and $a$symbolizes hyperparameters and is set gradually from 0.2 to 0.25. The main benefit of PRelu function is that the attained value is either a positive or negative integer. Hence the value comes in the negative axis won't be avoided as compared with other activation functions. Also, it acquires a function with a smaller slope with a change in the data distribution. Subsequently, the negative information is not lost, so the information loss of lung histopathological images is avoided. The detailed architect of proposed Dense AlexNet is described in Table 1.

Table 1: Dense AlexNet framework

Layer Name	***Size of the layer***	***Activations***
Input	256×256× 3	Batch normalisation
Convolution layer 1	64 filter, 11×11 pixel kernel size, 4 number of strides, 2 paddings and	PRelu and Batch normalization
	coefficient expansion of the dense layer	
Pooling layer 1	3×3 max pooling, 2 strides	-
Convolution layer 2	192 filter, 5×5 pixel kernel size, 5 number of strides, 2 padding	PRelu and Batch normalisation
Pooling layer 2	3×3 max pooling, 2 strides	-
Convolution layer 3	384 filter, 3×3 pixel kernel size, 3 number of strides, 1 padding	PRelu and Batch normalisation
Pooling layer 3	3×3 max pooling, 2 strides	-
Convolution layer 4	256 filter, 3×3 pixel kernel size, 3 number of strides, 1 padding	PRelu and Batch normalisation
Multi-scale convolution layer type	96 filter, 1×1 pixel kernel size, 1 number of stride, 1 padding	PRelu and Batch normalisation
	16 filter, 1×1 pixel kernel size, 1 number of stride, 1 padding	PRelu and Batch normalisation
	3×3 max pooling, 2 strides	-
	64 filter, 1×1 pixel kernel size, 1 number of stride, 1 padding	PRelu
	128 filter, 3×3 pixel kernel size, 1 number of stride, 1 padding	PRelu
	128 filter, 5×5 pixel kernel size, 1 number of stride, 1 padding	PRelu
	128 filter, 1×1 pixel kernel size, 1 number of stride, 1 padding	PRelu
	Concatenation	-
Convolution layer 5	256 filter, 3×3 pixel kernel size, 3 number of strides, 1 padding	PRelu and Batch normalisation
Fully connected layer 1	200 neuron nodes	PRelu and Dropout function
Fully connected layer 2	100 neuron nodes	PRelu and Dropout function
Fully connected layer 3	3 neuron nodes	PRelu and Dropout function

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

Output	1 neuron node	Soft max activation functions classification output

The stepwise procedure followed by the proposed Dense Alex Net framework is elaborated as follows:

1. The 1st layer comprises an input and batch normalization layer. The batch normalization layer carries the normalized input lung histopathological image. This layer can enhance the generalisation of the network model and improve the network convergence speed.
2. The 2ndlayer envelopes the convolution layer 1 by possessing 64 number of filters and a single11×11 kernel size of pixels. The dense convolution module is embedded in this layer for the improvement of extracting advanced lung disease features. Additionally, this layer involves PRelu as activation function for controlling the vanishing gradient problems. This layer also adopts the pooling layer where max pooling is utilised whereas the stride is 2, and the kernel size is fixed 3×3. When the pooling layer performance gets over, batch normalisation is incorporated.
3. The 3rdlayer comprises the convolution layer 2 by carrying 192 counts of filters with a 5×5kernel of pixel size and PRelu activation function and pooling layer 2. Aforementioned pooling layer involved here is max pooling with a kernel size of $3 \times 3$and stride allocated is 2. Followed by these layers, batch normalisation is performed.
4. The 4thlayer carries convolution layer 3 in which it consists of 384 filters by possessing a 3×3 kernel size of pixels. Same as the previous layers, PRelu activation function and max-pooling layer is utilised. Followed by these layers, batch normalisation is done.
5. The 5thlayer carries the convolution layer 4 where 256 filters is allocated with 3×3 kernel size of pixels. Additionally, this layer includes PRelu as activation function and adopts the batch normalisation phase.
6. The 6thlayer involves the multi-scale convolution phase. It is performed by carrying the number of filters from the topmost layer to the lowermost layer. Here the first layer of multi-scale convolution module is 96 and16 in which its kernel size allocated is 1×1 filter—the activation function involved in this layer id PRelu function. And the second layer of multi-scale convolution layer comprises 64, 128, 128and 128 with a kernel size of 1×1, 3×3,5×5 and 1×1 filters correspondingly. After the process gets over, the set of these multi-scale convolution layers is concatenated into a single layer.
7. The 7th layer employed in the proposed Dense AlexNet consists of convolution layer with module 5 where 256 filters are allotted. The kernel size used in this layer is 3×3 pixels. Same as the previous layers, PRelu activation function is employed. Followed by these layers, batch normalisation is applied over the network.
8. The 8th layer carries the first fully connected layer with 200 neurons and is then subjected to PRelu activation function. Followed by these layers, dropout function is performed.
9. The 9th layer incorporates the second fully connected layer by employing 100 neurons and is then handled by PReluactivation function. Followed by these layers, a dropout function is applied over the network model.
10. The final layer incorporates the last fully connected layer carrying three sets of neurons categories by employing the number of output lung disease classes.

Finally, the fully connected layer's obtained output is then subjected to the output layer for determining the classification of lung histopathological images. Here softmax activation function is performed by making the output classification rate as 1.0, and so the single output value is limited in the boundary range between 0 and1.

The pseudo-code of the proposed methodology is presented in the below table,

Table 2:Pseudo code of the proposed methodology

Input: lung histopathological image;

Output (C_m): Lung disease classification (ADC, SCC and benign);

Where ADC→Adeno Carcinoma, SCC→Squamous cell carcinoma;

//**Pre-processing procedure**//

- *Resize the input histopathological image;*
- *Smoothing using Gaussian filter using equation (1);*
- *Median value removal using equation (2);*
- *Original image removal using equation (3);*
- *Sharpen the obtained image using equation (4);*
- *Noise removal using median filter using equation (5);*

//**Contrast normalization **//

Perform Contrast Normalization using equation (6);

//**Augmentation **//

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

Do five-fold geometric transformation using equation (7), (8) and (9);
//**Recognition model: Dense AlexNet**//
Create Dense AlexNet structure for extracting advanced wider depth of features;
Fix the maximum number of iteration;
Set weights and bias functions;
Train CNN with its corresponding layers;
do
{
Procedure Predict CLASS (ADC, SCC and benign);
5 Convolution layer;
1 Multi scale convolution layer;
Batch Normalisation layer;
PRELU layer;
Max pooling layer;
3 Fully connected;
SoftmaxClassification Output;
}
return CLASS(ADC, SCC and benign);
end Procedure

The whole process is now capable of differentiating each lung disease. After the training process gets over, when new lung histopathological images are applied to the proposed methodology in the testing phase, lung disease type recognition is identified accurately. The overall accuracy of the proposed model shows the superiority of the proposed method.

IV. Experimental Result and Discussion

This section carries experimental outcomes of Dense Alex net-based histopathological lung image classification system. For lung cancer prediction and Classification, key steps incorporated are pre-processing, contrast normalization, and data augmentation before classification technique. The implementation was carried out in Python 3.7 processor -Intel i5 with 16GB RAM and 4GB graphics processor. For GPU analysis, the input images are tested on i3 processor with 4GB RAM and NVIDIA GeForce GTX 1650. The proposed method is validated, and tested results are compared with existing techniques to show that the proposed method performs better. Several performance measures like accuracy, precision, recall ,Roc Curve and F-measure are computed for analysing the performance of proposed research. Additionally, to justify the proposed algorithm's effectiveness, it is compared with the existing techniques such as Alex Net, Modified CNN and CNN.

4.1. Dataset Description

To show the validation of the proposed method, extensive experimentation is done on dataset [34] that contains 15,000 histopathological images with 3 classes. All images are in 768 X 768 pixels in size and are in JPEG file format. For each subject, images are generated from an original sample of HIPAA compliant and validated sources. There are three classes in the dataset, each with 5000 images. They are Lung benign tissue, Lung adenocarcinoma and Lung squamous cell carcinoma. To the best of our knowledge, it will be the most comprehensive data source currently available in the public domain and will facilitate further research efforts in this domain.

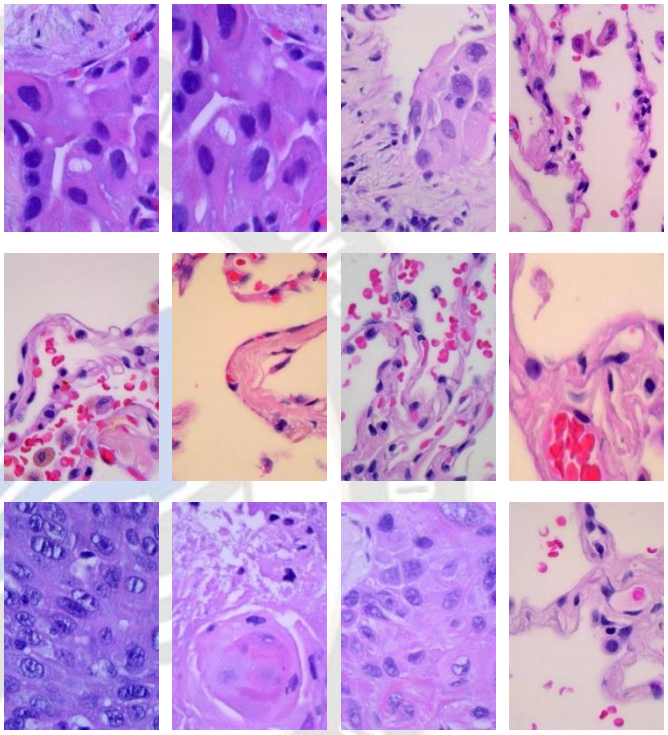


Figure 3: Sample input images

The above figure shows the sample input images taken to identify different types of lung diseases. The three lung tissue images are analysed using the proposed methodology. Before subjecting the input histopathological images into the proposed classification method, different steps are involved in the proposed approach to enhance system performance. The various steps involved are resizing, pre-processing, normalisation and augmentation. The outcome of lung images obtained by each stage of the proposed methodology is displayed in the following table.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

Table 3: Outcome of different lung tissue images

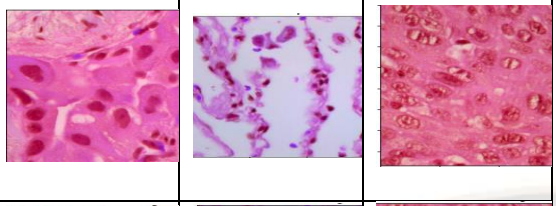
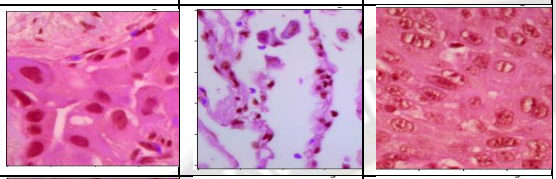
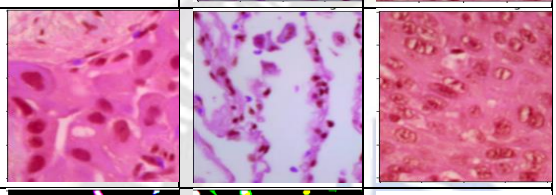
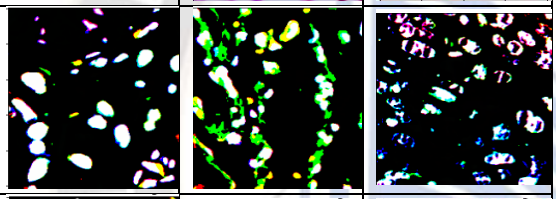
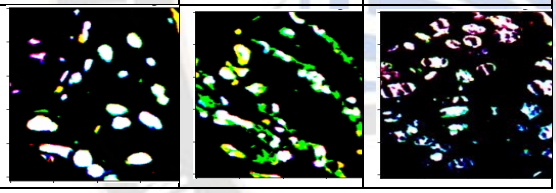
Image type	***Disease type***		
	Adenocarcinoma	Benign Tissue class	Squamous cell carcinoma
Input image			
Resized image			
Pre-processed image			
Normalised image			
Augmented image			

Table 3 shows the outcome of different histopathological lung tissue images like Adenocarcinoma, Benign Tissue class and Squamous cell carcinoma. The further section describes the evaluation metrics taken for analysis and offers the proposed and existing technique's outcome.

4.2. Evaluation metrics

To demonstrate the proposed system's effectiveness clearly, different test metrics like Accuracy, precision, F1_Score and Recall are evaluated. The mathematical expression is illustrated as follows.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (15)$$

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

$$Recall = \frac{TP}{TP+FN} \quad (17)$$

$$F1-Score = \frac{2\times(precision.recall)}{Precision+recall} \quad (18)$$

4.3. Performance analysis

The performance analysis of the proposed Dense AlexNet framework and existing techniques are compared in this section to show the effectiveness of the proposed methodology.

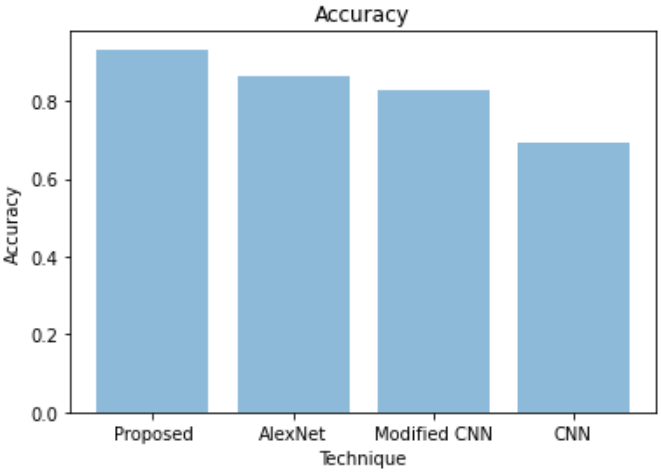
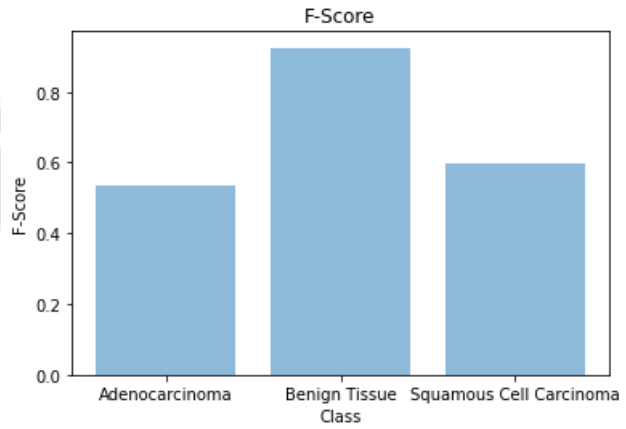


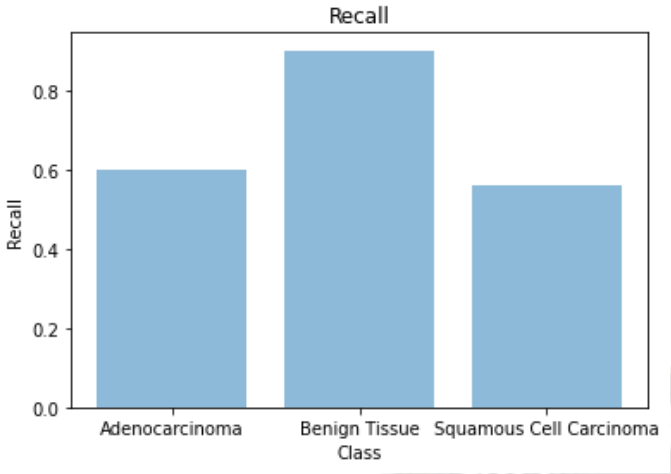
Figure 4: Accuracy comparison between proposed and existing techniques

From the figure4, it is understandable that the accuracy obtained for the proposed model is 93.24%. In contrast, the existing techniques like ALexNet, Modified CNN and CNN attained lesser accuracy measure, which is not up to the desired level for an appropriate diagnosis. Hence the proposed framework achieves higher classification measure than existing approaches. Likewise, other measures like F-measure, precision and recall are evaluated for the lung cancer classification model, and it's graphically shown in the below graphs. It shows that different types of lung cancer tissues obtained with their corresponding measures are provided.



(a)

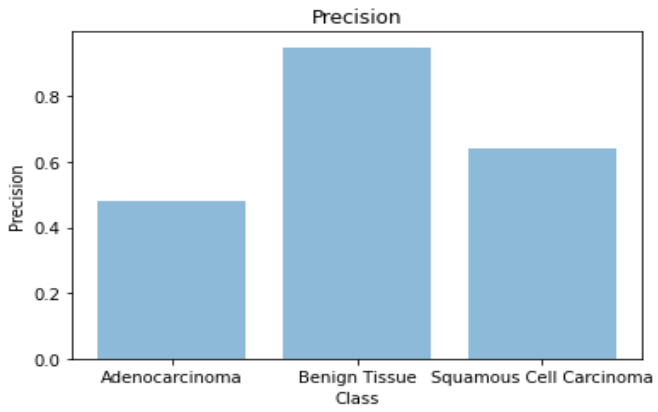
International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*



(c)

Figure 5. Obtained results between different ypes of lung cancer tissue vs (a)F1_Score, (b) Precision and (c)Recall

The above figure 5 represents the number of classes correctly identified for the respective measures taken. The three classes taken for lung disease classification analysis is Adenocarcinoma, Benign Tissue class and Squamous cell carcinoma. The correctly determined measures are computed for these multiple lung disease classes and shown in the above-plotted graph. Almost all the values attained are higher for the respective classes. The proposed scheme achieves the maximum outcome for all the performance metrics compared with the lung disease



(b)

classification. Therefore from the overall analysis, it is clear that the anticipated method has given a better performance than the conventional techniques.

V. Conclusion

The proposed technique introduces a novel approach named Dense AlexNet for the classification of input histopathological lung images. The proposed method's chief objective is to categorise three stages of lung disease efficiently. The pre-processing step carries several steps to eliminate noisy contents contained in the input image. Next step is to perform contrast normalisation phase, which is done for maintaining the same illumination factor. This helps to make the network model to learn faster. Then data augmentation step is carried out to avoid over-fitting problems occurring in the network. Finally, classification is performed using the proposed Dense AlexNet model. This, in turn, improves classification accuracy by automatically achieving a more comprehensive range of features. Thus the performance of the proposed method gets enhanced by reducing the computational complexity. The proposed methodology achieved efficient outcome compared with the existing approaches in terms of accuracy, F-measure, precision, and recall.

Reference

[1]. González Maldonado, S., Motsch, E., Trotter, A., Kauczor, H.U., Heussel, C.P., Hermann, S., Zeissig, S.R., Delorme, S. and Kaaks, R., 2020. Overdiagnosis in lung cancer screening: Estimates from the German Lung Cancer Screening Intervention Trial. International Journal of Cancer.

[2]. Ying, L., Du, L., Zou, R., Shi, L., Zhang, N., Jin, J., Xu, C., Zhang, F., Zhu, C., Wu, J. and Chen, K., 2020. Development of a serum miRNA panel for detection of early stage non-small cell lung cancer. Proceedings of the National Academy of Sciences, 117(40), pp.25036-25042.

[3]. Skoulidis, F. and Heymach, J.V., 2019. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. Nature reviews Cancer, 19(9), pp.495-509.

[4]. Ardila, D., Kiraly, A.P., Bharadwaj, S., Choi, B., Reicher, J.J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G. and Naidich, D.P., 2019. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature medicine, 25(6), pp.954-961.

[5]. Matsubara, T., Ochiai, T., Hayashida, M., Akutsu, T. and Nacher, J.C., 2019. Convolutional neural network approach to lung cancer classification integrating protein interaction network and gene expression profiles. Journal of bioinformatics and computational biology, 17(03), p.1940007.

[6]. Zhang, C., Sun, X., Dang, K., Li, K., Guo, X.W., Chang, J., Yu, Z.Q., Huang, F.Y., Wu, Y.S., Liang, Z. and Liu, Z.Y., 2019. Toward an expert level of lung cancer detection and classification using a deep convolutional neural network. The oncologist, 24(9), p.1159.

[7]. Linning, E., Lu, L., Li, L., Yang, H., Schwartz, L.H. and Zhao, B., 2019. Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. Academic radiology, 26(9), pp.1245-1252.

[8]. Ahmed, S.R.A., Al Barazanchi, I., Mhana, A. and Abdulshaheed, H.R., 2019. Lung cancer classification using data mining and supervised learning algorithms on multi-dimensional data set. Periodicals of Engineering and Natural Sciences, 7(2), pp.438-447.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 2
DOI: https://doi.org/10.17762/ijritcc.v11i2.6133
***Article Received**: 24 December 2022 **Revised**: 10 January 2023 **Accepted**: 20 January 2023*

[9]. Zhang, M., He, J., Li, T., Hu, H., Li, X., Xing, H., Wang, J., Yang, F., Ma, Q., Liu, B. and Tang, C., 2019. Accurate classification of non-small cell lung cancer (NSCLC) pathology and mapping of EGFR mutation spatial distribution by ambient mass spectrometry imaging. Frontiers in oncology, 9, p.804.

[10]. Zhang, M., He, J., Li, T., Hu, H., Li, X., Xing, H., Wang, J., Yang, F., Ma, Q., Liu, B. and Tang, C., 2019. Accurate classification of non-small cell lung cancer (NSCLC) pathology and mapping of EGFR mutation spatial distribution by ambient mass spectrometry imaging. Frontiers in oncology, 9, p.804.

[11]. Sang, J., Alam, M.S. and Xiang, H., 2019, May. Automated detection and classification for early stage lung cancer on CT images using deep learning. In Pattern Recognition and Tracking XXX (Vol. 10995, p. 109950S). International Society for Optics and Photonics.

[12]. Palani, D. and Venkatalakshmi, K., 2019. An IoT based predictive modelling for predicting lung cancer using fuzzy cluster based segmentation and classification. Journal of medical systems, 43(2), p.21.

[13]. Shakeel, P.M., Burhanuddin, M.A. and Desa, M.I., 2019. Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. Measurement, 145, pp.702-712.

[14]. Wang, X., Chen, H., Gan, C., Lin, H., Dou, Q., Tsougenis, E., Huang, Q., Cai, M. and Heng, P.A., 2019. Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. IEEE Transactions on Cybernetics.

[15]. ALzubi, J.A., Bharathikannan, B., Tanwar, S., Manikandan, R., Khanna, A. and Thaventhiran, C., 2019. Boosted neural network ensemble classification for lung cancer disease diagnosis. Applied Soft Computing, 80, pp.579-591.

[16]. Bhuvaneswari, C., Aruna, P. and Loganathan, D., 2014. A new fusion model for classification of the lung diseases using genetic algorithm. Egyptian Informatics Journal, 15(2), pp.69-77.

[17]. Ma, L., Liu, X., Song, L., Zhou, C., Zhao, X. and Zhao, Y., 2015. A new classifier fusion method based on historical and on-line classification reliability for recognising common CT imaging signs of lung diseases. Computerized Medical Imaging and Graphics, 40, pp.39-48.

[18]. Karthiga, B. and Rekha, M., 2020. Feature extraction and I-NB classification of CT images for early lung cancer detection. Materials Today: Proceedings.

[19]. Suresh, S. and Mohan, S., 2019. NROI based feature learning for automated tumor stage classification of pulmonary lung nodules using deep convolutional neural networks. Journal of King Saud University-Computer and Information Sciences.

[20]. Joyseeree, R., Otálora, S., Müller, H. and Depeursinge, A., 2019. Fusing learned representations from Riesz Filters and Deep CNN for lung tissue classification. Medical image analysis, 56, pp.172-183.

[21]. Wu, J., Tan, Y., Chen, Z. and Zhao, M., 2018. Decision based on big data research for non-small cell lung cancer in medical artificial system in developing country. Computer methods and programs in biomedicine, 159, pp.87-101.

[22]. Chang, Y., Kim, N., Lee, Y., Lim, J., Seo, J.B. and Lee, Y.K., 2012. Fast and efficient lung disease classification using hierarchical one-against-all support vector machine and cost-sensitive feature selection. Computers in biology and medicine, 42(12), pp.1157-1164.

[23]. Savitha, G. and Jidesh, P., 2020. A holistic deep learning approach for identification and classification of sub-solid lung nodules in computed tomographic scans. Computers & Electrical Engineering, 84, p.106626.

[24]. Li, J., Wang, Y., Song, X. and Xiao, H., 2018. Adaptive multinomial regression with overlapping groups for multi-class classification of lung cancer. Computers in biology and medicine, 100, pp.1-9.

[25]. Lakshmanaprabu, S.K., Mohanty, S.N., Shankar, K., Arunkumar, N. and Ramirez, G., 2019. Optimal deep learning model for classification of lung cancer on CT images. Future Generation Computer Systems, 92, pp.374-382.

[26]. Liu, Y., Wang, H., Gu, Y. and Lv, X., 2019. Image classification toward lung cancer recognition by learning deep quality model. Journal of Visual Communication and Image Representation, 63, p.102570.

[27]. Teramoto, A., Yamada, A., Kiriyama, Y., Tsukamoto, T., Yan, K., Zhang, L., Imaizumi, K., Saito, K. and Fujita, H., 2019. Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network. Informatics in Medicine Unlocked, 16, p.100205.

[28]. Liu, X., Ma, L., Song, L., Zhao, Y., Zhao, X. and Zhou, C., 2015. Recognising common CT imaging signs of lung diseases through a new feature selection method based on Fisher criterion and genetic optimisation. IEEE J. Biomed. Health Informatics, 19(2), pp.635-647.

[29]. Hawkins, S.H., Korecki, J.N., Balagurunathan, Y., Gu, Y., Kumar, V., Basu, S., Hall, L.O., Goldgof, D.B., Gatenby, R.A. and Gillies, R.J., 2014. Predicting outcomes of nonsmall cell lung cancer using CT image features. IEEE access, 2, pp.1418-1426.

[30]. Petousis, P., Winter, A., Speier, W., Aberle, D.R., Hsu, W. and Bui, A.A., 2019. Using Sequential Decision Making to Improve Lung Cancer Screening Performance. IEEE Access, 7, pp.119403-119419.

[31]. Tan, T., Li, Z., Liu, H., Zanjani, F.G., Ouyang, Q., Tang, Y., Hu, Z. and Li, Q., 2018. Optimise transfer learning for lung diseases in bronchoscopy using a new concept: sequential fine-tuning. IEEE journal of translational engineering in health and medicine, 6, pp.1-8.

[32]. https://www.kaggle.com/andrewmvd/lung-and-colon-cancer-histopathological-images.